

Figure 1: Nucleotide Sequence of TbF14  
Sheet 1 of 4

FEATURES	Location/Qualifiers
misc_feature	5072..5095
	/note="His tag coding region"
misc_feature	5096..7315
	/note="MtB81 coding region"
misc_feature	7316..8594
	/note="Mo2 coding region"

TGGCGAATGGGACGCGCCCTGTAGCGGCGCATTAAAGCGCGGGTGTGGTGGTTACGCGCAGCGT  
GACCGCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTTCGCTTTCTTCCCTTCTTTCTCGCCAC  
GTTCCGCCGGCTTTCCCCGTCAAGCTCTAAATCGGGGGCTCCCTTTAGGGTTCGGATTTAGTGCTTT  
ACGGCACCTCGACCCCAAAAACTTGATTAGGGTGATGGTTCACGTAGTGGGCCATCGCCCTGATA  
GACGGTTTTTCGCCCTTTGACGTTGGAGTCCACGTTCTTTAATAGTGGACTCTTGTTCCAACTGG  
AACAACTCAACCCTATCTCGGTCTATTCTTTTGATTATAAGGGATTTTGCCGATTTCCGCCCTA  
TTGGTTAAAAATGAGCTGATTTAACAAAAATTTAACCGGAATTTTAACAAAATATTAACGTTTAC  
AATTTACAGGTGGCACTTTTCGGGGAATGTGCGCGGAACCCCTATTTGTTTATTTTCTAAATACA  
TTCAAATATGTATCCGCTCATGAATTAATTCTTAGAAAACTCATCGAGCATCAAATGAACTGCA  
ATTTATTCATATCAGGATTATCAATACCATATTTTGAAGAGCCGTTTCTGTAATGAAGGAGAAA  
ACTCACCGAGGCAGTTCATAGGATGGCAAGATCCTGGTATCGGTCTGCGATTCCGACTCGTCCAA  
CATCAATACAACCTATTAATTTCCCTCGTCAAAAATAAGGTTATCAAGTGAGAAATCACCATGAG  
TGACGACTGAATCCGGTGAGAATGGCAAAAGTTTATGCATTTCTTTCCAGACTTGTTCAACAGGCC  
AGCCATTACGCTCGTCATCAAAATCACTCGCATCAACCAACCGTTATTCATTCTGTGATTGCGCCT  
GAGCGAGACGAAATACGCGATCGCTGTTAAAGGACAATTACAAACAGGAATCGAATGCAACCGGC  
GCAGGAACACTGCCAGCGCATCAACAATATTTTACCTGAATCAGGATATTCTTCTAATACCTGGA  
ATGCTGTTTTCCCGGGGATCGCAGTGGTGAGTAACCATGCATCATCAGGAGTACGGATAAAATGCT  
TGATGGTCGGAAGAGGCATAAATCCGTCAGCCAGTTTAGTCTGACCATCTCATCTGTAACATCAT  
TGGCAACGCTACCTTTGCCATGTTTCAGAAACAACCTCTGGCGCATCGGGCTTCCCATACAATCGAT  
AGATTGTGCGACCTGATTGCCCCGACATTATCGCGAGCCCATTTATACCCATATAAATCAGCATCCA  
TGTTGGAATTTAATCGCGGCCTAGAGCAAGACGTTTCCCGTTGAATATGGCTCATAACACCCCTTG  
TATTACTGTTTATGTAAGCAGACAGTTTTATTGTTTCATGACCAAAATCCCTTAACGTGAGTTTTCG  
TTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTTCTGCGC  
GTAATCTGCTGCTTGCAACAAAAAACCACCGCTACCAGCGGTGGTTTGTGTTGCCGGATCAAGAG  
CTACCAACTCTTTTCCGAAGGTAACCTGGCTTCAGCAGAGCGCAGATACCAAATACTGTCTTCTA  
GTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGCTA  
ATCCTGTTACCAAGTGGCTGCTGCCAGTGGCGATAAGTCTGTGCTTACCGGGTTGGACTCAAGACGA  
TAGTTACCGGATAAGGCGCAGCGGTGGGGCTGAACGGGGGGTTCGTGCACACAGCCAGCTTGGAG  
CGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAA  
GGGAGAAAGGCGGACAGGTATCCGTAAGCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTT  
CCAGGGGGAACGCCTGGTATCTTTATAGTCTGTGCGGTTTCGCCACCTCTGACTTGAGCGTCTGA  
TTTTTGATGCTCGTCAGGGGGCGGAGCCTATGGAAAAACGCCAGCAACGCGGCCTTTTTACGG  
TTCCTGGCCTTTTGCTGGCCTTTTGCTCACATGTTCTTCTGCGTTATCCCTGATTCTGTGGAT  
AACCGTATTACCGCCTTTGAGTGAGCTGATACCGCTCGCCGACCGCAACGACCGAGCGCAGCGAG  
TCAGTGAGCGAGGAAGCGGAAGAGCGCCTGATGCGGTATTTCTCCTTACGCATCTGTGCGGTATT  
TCACACCGCATATATGGTGCACTCTCAGTACAATCTGCTCTGATGCCGCATAGTTAAGCCAGTATA

09688672-101000

Figure 1: Nucleotide Sequence of TbF14  
Sheet 2 of 4

CACTCCGCTATCGCTACGTGACTGGGTTCATGGCTGCGCCCCGACACCCGCCAACACCCGCTGACGC  
GCCCTGACGGGCTTGTCTGCTCCCGGCATCCGCTTACAGACAAGCTGTGACCGTCTCCGGGAGCTG  
CATGTGTCAGAGGTTTTACCGTCATCACCGAAACGCGCGAGGCAGCTGCGGTAAAGCTCATCAGC  
GTGGTCGTGAAGCGATTACAGATGTCTGCCTGTTTCATCCGCGTCCAGCTCGTTGAGTTTCTCCAG  
AAGCGTTAATGTCTGGCTTCTGATAAAGCGGGCCATGTTAAGGGCGGTTTTTCTGTTTGGTTCAC  
TGATGCCTCCGTGTAAGGGGGATTCTGTTTCATGGGGGTAATGATACCGATGAAACGAGAGAGGAT  
GCTCACGATACGGGTTACTGATGATGAACATGCCCGGTTACTGGAACGTTGTGAGGGTAAACAACT  
GGCGGTATGGATGCGGCGGGACCAGAGAAAAATCACTCAGGGTCAATGCCAGCGCTTCGTTAATAC  
AGATGTAGGTGTTCCACAGGGTAGCCAGCAGCATCCTGCGATGCAGATCCGGAACATAATGGTGCA  
GGGCGCTGACTTCCGCGTTTCCAGACTTTACGAAACACGGAAACCGAAGACCATTTCATGTTGTTGC  
TCAGGTGCGCAGACGTTTTTGCAGCAGCAGTCGCTTCACGTTTCGCTCGCTATCGGTGATTTCATTCTG  
CTAACCAGTAAGGCAACCCCGCCAGCCTAGCCGGGTCTCTCAACGACAGGAGCACGATCATGCGCAC  
CCGTGGGGCCGCCATGCCGGCGATAATGGCCTGCTTCTCGCCGAAACGTTTGGTGGCGGGACCAGT  
GACGAAGGCTTGAGCGAGGGCGTGCAAGATTCCGAATACCGCAAGCGACAGGCCGATCATCGTCGC  
GCTCCAGCGAAAGCGGTCTCTCGCCGAAAATGACCCAGAGCGCTGCCGGCACCTGTCTACGAGTTG  
CATGATAAAGAAGACAGTCATAAGTGCGGCGACGATAGTCATGCCCCGCGCCACCAGGAGGAGCT  
GACTGGGTTGAAGGCTCTCAAGGGCATCGGTTCGAGATCCCGGTGCCTAATGAGTGAGCTAACTTAC  
ATTAATTGCGTTGCGCTCACTGCCCCGTTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATG  
AATCGGCCAACGCGCGGGGAGAGGCGGTTTTGCGTATTGGGCGCCAGGGTGGTTTTTCTTTTACCA  
GTGAGACGGGCAACAGCTGATTGCCCTTACCCGCTGGCCCTGAGAGAGTTGCAGCAAGCGGTCCA  
CGCTGGTTTGCCCCAGCAGGCGAAAATCCTGTTTGATGGTGGTTAACGGCGGGATATAACATGAGC  
TGTCTTCGGTATCGTCGTATCCCACTACCGAGATATCCGCACCAACGCGCAGCCCGGACTCGGTAA  
TGGCGCGCATTGCGCCCAGCGCCATCTGATCGTTGGCAACCAGCATCGCAGTGGGAACGATGCCCT  
CATTAGCATTTGCATGGTTTGTGAAAACCGGACATGGCACTCCAGTCGCCTTCCCGTTCCGCTA  
TCGGCTGAATTTGATTGCGAGTGAGATATTTATGCCAGCCAGCCAGACGCGCGCCGAGACAG  
AACTTAATGGGCCCCGCTAACAGCGCGATTTGCTGGTGACCCAATGCGACCAGATGCTCCACGCCCA  
GTCGCGTACCGTCTTCATGGGAGAAAAATAACTGTTGATGGGTGTCTGGTCAGAGACATCAAGAA  
ATAACGCCGGAACATTAGTGAGGAGCTTCCACAGCAATGGCATCCTGGTCATCCAGCGGATAGT  
TAATGATCAGCCCACTGACGCGTTGCGCGAGAAGATTGTGCACCGCGCTTTACAGGCTTCGACGC  
CGCTTCGTTCTACCATCGACACCACCGCTGGCACCCAGTTGATCGGCGCGAGATTTAATCGCCC  
CGACAATTTGCGACGGCGCGTGCAGGGCCAGACTGGAGGTGGCAACGCCAATCAGCAACGACTGTT  
TGCCCGCCAGTTGTTGTGCCACGCGGTGGGAATGTAATTCAGCTCCGCCATCGCCGCTTCCACTT  
TTTCCCGCGTTTTTCGAGAAACGTGGCTGGCCTGGTTTACCACGCGGGAACGGTCTGATAAGAGA  
CACCGGCATACTCTGCGACATCGTATAACGTTACTGGTTTACATTACCACCCTGAATTGACTCT  
CTTCCGGGCGCTATCATGCCATACCGCGAAAGGTTTTCGCCATTTCGATGGTGTCCGGGATCTCGA  
CGCTCTCCCTTATGCGACTCCTGCATTAGGAAGCAGCCCAGTAGTAGTTGAGGCCGTTGAGCACC  
GCCGCCGCAAGGAATGGTGCATGCAAGGAGATGGCGCCCAACAGTCCCCCGGCCACGGGGCCTGCC  
ACCATACCCACGCCGAAACAAGCGCTCATGAGCCCGAAGTGGCGAGCCCGATCTTCCCCATCGGTG  
ATGTCGGCGATATAGGCGCCAGCAACCGCACCTGTGGCGCCGGTGTGCGGCCACGATGCGTCCG  
GCGTAGAGGATCGAGATCTCGATCCCGCGAAATTAATACGACTCACTATAGGGGAATTGTGAGCGG  
ATAACAATTTCCCTCTAGAAATAATTTGTTTAACTTTAAGAAGGAGATATACATATGCAGCATCA  
CCACCATCACCACACTGATCGCGTGTGCGTGGGCAACTTGCATCGCTCGGGTGTCTACGACTT  
CGTGAACAATGAAGCCCTGCCTGGCACCGATATCGACCCGACAGCTTCTGGGCGGGCGTTCGACAA

09688672-171777

Figure 1: Nucleotide Sequence of TbF14  
Sheet 3 of 4

GTCTGCGCCGACCTGACCCCGCAGAACCAAGCTCTGTTGAACGCCCCGCGACGAGCTGCAGGCGCAG  
ATCGACAAGTGGCACC GGCGTCGGGTGATCGAGCCCATCGACATGGATGCCTACCGCCAGTTCCTC  
ACCGAGATCGGCTACCTGCTTCCCGAACCTGATGACTTCACCATCACCACGTCCGGTGTGACGCT  
GAGATCACCACGACCGCCGGCCCCCAGCTGGTGGTGCCGGTGCTCAACGCGCGGTTTGTCTGTAAC  
GCGGCCAACGCTCGCTGGGGCTCCCTCTACGACGCCTTGTATGGCACCAGTGTATCCCCGAGACC  
GACGGCGCCGAAAAAGGCCCCACGTACAACAAGGTTCTGTGGCGACAAGGTGATCGCGTATGCCCGC  
AAGTTCCTCGACGACAGTGTTCGCTGTCTGTCGGGTTCTTTGGCGACGCCACCGGTTTCACAGTG  
CAGGATGGCCAGCTCGTGGTTGCCTTGCCGGATAAGTCCACCGGCCTGGCCAACCCCGGCCAGTTC  
GCCGGCTACACCGGCGCAGCCGAGTCGCCGACATCGGTGCTGCTAATCAATCACGGTTTGCACATC  
GAGATCCTGATCGATCCGGAGTCGCAGGTCCGCACCACCGACCGGGCCGGCGTCAAGGACGTGATC  
CTGGAATCCGCGATCACCACGATCATGGACTTCGAGGACTCGGTGGCCGCGGTGGACGCCGCCGAC  
AAGGTGCTGGGTTATCGGAACTGGCTCGGCCTGAACAAGGGCGACCTGGCAGCAGCGGTAGACAAG  
GACGGCACC GCTTTCTGCGGGTGCTCAATAGGGACCGGAACTACACCGCACCCGGCGGTGGCCAG  
TTCACGCTGCCTGGACGCAGCCTCATGTTCTGTCGCAACGTCCGTCACTTGATGACGAATGACGCC  
ATCGTCGACACTGACGGCAGCGAGGTGTTTGAAGGCATCATGGATGCCCTATTACCGGCCTGATC  
GCCATCCACGGGCTAAAGGCCAGCGACGTCAACGGGCGGCTGATCAACAGCCGCACCGGCTCCATC  
TACATCGTCAAGCCGAAGATGCACGGTCCGGCCGAGGTGGCGTTTACCTGCGAACTGTTACGCCG  
GTTGAAGATGTGCTGGGGTTGCCGCAAAACACCATGAAGATCGGCATCATGGACGAGGAACGCCGG  
ACCACGGTCAACCTCAAGGCGTGCATCAAAGCTGCCGCGGACCGCGTGGTGTTCATCAACACCGGG  
TTCCTGGACCGCACCGGCGATGAAATCCACACCTCGATGGAGGCCGGCCCGATGGTGC GCAAGGGC  
ACCATGAAGAGCCAGCCGTGGATCTTGGCCTACGAGGACCACAACGTTCGATGCCGGCCTGGCCGCC  
GGGTTACGCGGCCGAGCCAGGTCCGGCAAGGGCATGTGGACAATGACCGAGCTGATGGCCGACATG  
GTCGAGACAAAAATCGCCAGCCGCGCGCCGGGGCCAGCACCGCCTGGGTTCCTCTCCCACTGCG  
GCCACCCTGCATGCGCTGCACTACCACCAGGTTCGACGTGCGCGCGGTGCAACAAGGACTGGCGGGG  
AAGCGTCGCGCCACCATCGAACAATTGCTGACCATTCCGCTGGCCAAGGAATTGGCCTGGGCTCCC  
GACGAGATCCGCGAAGAGGTTCGACAACAATGTCAATCCATCCTCGGCTACGTGGTTCTGCTGGGTT  
GATCAAGGTGTCCGCTGCTCGAAGGTGCCCAGCATCCACGACGTGCGCGCTCATGGAGGACCGGGCC  
ACGCTGCGAATCTCCAGCCAATTGTTGGCCAATGGCTGCGCCACGGTGTGATCACCAGCGCGGAT  
GTGCGGGCCAGCTTGGAGCGGATGGCGCCGTTGGTTCGATCGACAAAACGCGGGCGACGTGGCATA  
CGACCGATGGCACCCAACTTCGACGACAGTATCGCCTTCCTGGCCGCGCAGGAGCTGATCTTGTC  
GGGGCCCAGCAGCCCAACGGCTACACCGAGCCGATCCTGCACCGACGTCGTCCGGAGTTTAAGGCC  
CGGGCCGCTGAGAAGCCGGCCCCATCGGACAGGGCCGGTGACGATGCGGCCAGGGTGCAGAAGTAC  
GGCGGATCCTCGGTGGCCGACGCCGAACGGATTTCGCCGCGTCGCCGAACGCATCGTCGCCACCAAG  
AAGCAAGGCAATGACGTCGTCTGCTGCTCTGCTGATGGGGGATACACCGACGACCTGCTGGAT  
CTGGCTCAGCAGGTGTGCCCGGCGCCGCGCCTCGGGAGCTGGACATGCTGCTTACCGCCGGTGAA  
CGCATCTCGAATGCGTTGGTGGCCATGGCCATCGAGTCGCTCGGCGCGCATGCCCGGTCTGTTACC  
GGTTCGAGGCCGGGGTGATCACCACCGGCACCCACGGCAACGCCAAGATCATCGACGTACGCCG  
GGGCGGCTGCAAACCGCCCTTGAGGAGGGGCGGGTCTGTTTGGTGGCCGGATTCCAAGGGGTGAGC  
CAGGACACCAAGGATGTACGACGTTGGGCCGCGCGGCTCGGACACCACCGCCGTCGCCATGGCC  
GCCGCGCTGGGTGCCGATGTCTGTGAGATCTACACCGACGTGGACGGCATCTTCAGCGCCGACCCG  
CGCATCGTGC GCAACGCCCGAAAGCTCGACACCGTGACCTTCGAGGAAATGCTCGAGATGGCGGCC  
TGCGGCGCCAAGGTGCTGATGCTGCGTGGGATACGCTCGCCGCCATAATATTCCGGTGCAC  
GTCCGGTCTGCTACTCGGACAGACCGGGCACCGTCTGTTGTCGGATCGATCAAGGACGTACCCATG

096857.101000

Figure 1: Nucleotide Sequence of TbF14

Sheet 4 of 4

GAAGACCCCATCCTGACCGGAGTCGCGCACGACCGCAGCGAGGCCAAGGTGACCATCGTCGGGCTG  
CCCCACATCCCCGGGTATGCGGCCAAGGTGTTTAGGGCGGTGGCCAGACGCCGACGTCAACATCGA  
CATGGTGCTGCAGAACGTCTCCAAGGTCGAGGACGGCAAGACCGACATCACCTTACCTGCTCCCG  
CAGACGTCGGGCCCGCCGTGGAAAACTGGACTCGCTCAGAAACGAGATCGGCTTCTACACAG  
CTGCTGTACGACGACCACATCGGCAAGGTATCGCTGATCGGTGCCGGCATGCGCAGCCACCCCGG  
GTCACCGCGACGTTCTGTGAGGCGCTGGCGCGGTGGGGGTCAACATCGAGCTGATCTCCACCTCG  
GAAGATCAGAGATCTCGGTGTTGTGCCGCGACACCGAAGGTGACGCGGGGACGGGACGGTAGATGGCC  
AAGCGTTCGGGCTCGGCGCGACGAGGAGCCACGGTGTACGCGGGCGGCTCGAGCACCACCAC  
TGTCATAAGTGAATTCATCGATGTGCAGATATCCATCACACTGGCGGGCGGCTGCCACCGCTG  
ACCACCACTGAGATCCGGCTGCTAACAAAGCCCGAAAGGAGCTGAGTTGGCTGCTGCCACCGCTG  
AGCAATAACTAGCATAACCCCTTGGGGCCTCTAAACGGGTCTTGAGGGTTTTTTGCTGAAAGGAG  
GAACTATATCCGGAT

**THE UNIVERSITY OF CHICAGO**

Figure 2: Nucleotide sequence of TbF15  
Sheet 1 of 4

FEATURES	Location/Qualifiers
misc_feature	5072..5095
	/note="His tag coding region"
misc_feature	5096..5293
	/note="Ra3 coding region"
misc_feature	5294..6346
	/note="38kD coding region"
misc_feature	6347..6643
	/note="38-1 coding region"
misc_feature	6644..8023
	/note="FL TbH4 coding region"

TGGCGAATGGGACGCGCCCTGTAGCGGCGCATTAAAGCGCGGGCGGGTGTGGTGGTTACGCGCAGCGT  
GACCGCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTTCGCTTTCTTCCCTTCTTTCTCGCCAC  
GTTTCGCCGGCTTTCCCCGTCAAGCTCTAAATCGGGGGCTCCCTTTAGGGTTCGGATTTAGTGCTTT  
ACGGCACCTCGACCCCAAAAACTTGATTAGGGTGATGGTTCACGTAGTGGGCCATCGCCCTGATA  
GACGGTTTTTCGCCCTTTGACGTTGGAGTCCACGTTCTTTAATAGTGGACTCTTGTTCCAACTGG  
AACAACTCAACCCTATCTCGGTCTATTCTTTTGATTATAAGGGATTTTGCCGATTTCCGGCCTA  
TTGGTTAAAAAATGAGCTGATTTAACAAAAATTTAACGCGAATTTTAACAAAAATTTAACGTTTAC  
AATTTGAGGTGGCACTTTTCGGGGAAATGTGCGCGGAACCCCTATTTGTTTATTTTTCTAAATACA  
TTCAAATATGTATCCGCTCATGAATTAATTCTTAGAAAACTCATCGAGCATCAAATGAACTGCA  
ATTTATTCATATCAGGATTATCAATACCATATTTTGAAGAGCCGTTTCTGTAATGAAGGAGAAA  
ACTCACCAGGCAGTTCATAGGATGGCAAGATCCTGGTATCGGTCTGCGATTCCGACTCGTCCAA  
CATCAATACAACCTATTAATTTCCCTCGTCAAAAAATAAGGTTATCAAGTGAGAAATCACCATGAG  
TGACGACTGAATCCGGTGAGAATGGCAAAAGTTTATGCATTTCTTTCCAGACTTGTTCAACAGGCC  
AGCCATTACGCTCGTCATCAAAATCACTCGCATCAACCAACCGTTATTCATTCTGTGATTGCGCCT  
GAGCGAGACGAAATACGCGATCGCTGTTAAAAGGACAATTACAAACAGGAATCGAATGCAACCGGC  
GCAGGAACACTGCCAGCGCATCAACAATATTTTACCTGAATCAGGATATTCTTCTAATACCTGGA  
ATGCTGTTTTTCCCGGGGATCGCAGTGGTGAGTAACCATGCATCATCAGGAGTACGGATAAAATGCT  
TGATGGTTCGGAAGAGGCATAAATTCGCTCAGCCAGTTTAGTCTGACCATCTCATCTGTAACATCAT  
TGGCAACGCTACCTTTGCCATGTTTCAGAAACAACTCTGGCGCATCGGGCTTCCCATACAATCGAT  
AGATTGTCGCACCTGATTGCCCCGACATTATCGCGAGCCCATTTATACCCATATAAATCAGCATCCA  
TGTTGGAATTTAATCGCGGCCTAGAGCAAGACGTTTCCCGTTGAATATGGCTCATAACACCCCTTG  
TATTACTGTTTATGTAAGCAGACAGTTTTATTGTTTCATGACCAAAATCCCTTAACGTGAGTTTTCG  
TTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTTTCTGCGC  
GTAATCTGCTGCTTGCAAACAAAAAACACCGCTACCAGCGGTGGTTTGGTTTGGCCGATCAAGAG  
CTACCAACTCTTTTTCCGAAGGTAAGTGGCTTCAGCAGAGCGCAGATAACCAATACTGTCCTTCTA  
GTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGCTA  
ATCCTGTTACCAAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTGGACTCAAGACGA  
TAGTTACCGGATAAGGCGCAGCGGTGCGGCTGAACGGGGGGTTCGTGCACACAGCCCAGCTTGGAG  
CGAACGACCTACACCGAAGTGAATACCTACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAA  
GGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCAGGGTTCGGAACAGGAGAGCGCACGAGGGAGCTT  
CCAGGGGGAAACGCCTGGTATCTTTATAGTCCTGTGCGGTTTCGCCACCTCTGACTTGAGCGTCCA  
TTTTTGTGATGCTCGTCAGGGGGGCGGAGCCTATGGAAAAACGCCAGCAACGCGGCCTTTTTACGG

09588672-11111111

Figure 2: Nucleotide sequence of TbF15  
Sheet 2 of 4

TTCCTGGCCTTTTGCTGGCCTTTTGCTCACATGTTCTTTCCTGCGTTATCCCCTGATTCTGTGGAT  
AACCGTATTACCGCCTTTGAGTGAGCTGATACCGCTCGCCGCAGCCGAACGACCGAGCGCAGCGAG  
TCAGTGAGCGAGGAAGCGGAAGAGCGCCTGATGCGGTATTTTCTCCTTACGCATCTGTGCGGTATT  
TCACACCGCATATATGGTGCCTCTCAGTACAATCTGCTCTGATGCCGCATAGTTAAGCCAGTATA  
CACTCCGCTATCGCTACGTGACTGGGTATGGCTGCGCCCCGACACCCGCCAACACCCGCTGACGC  
GCCCTGACGGGCTTGTCTGCTCCCGGCATCCGCTTACAGACAAGCTGTGACCGTCTCCGGGAGCTG  
CATGTGTGAGAGGTTTTACCGTCATCACCGAAACGCGCGAGGCAGCTGCGGTAAAGCTCATCAGC  
GTGGTCGTGAAGCGATTACAGATGTCTGCCTGTTTATCCGCGTCCAGCTCGTTGAGTTTCTCCAG  
AAGCGTTAATGTCTGGCTTCTGATAAAGCGGGCCATGTTAAGGGCGGTTTTTCTCTGTTTGGTCAC  
TGATGCCTCCGTGTAAGGGGGATTCTGTTCATGGGGTAATGATACCGATGAAACGAGAGAGGAT  
GCTCACGATACGGGTTACTGATGATGAACATGCCCGGTTACTGGAACGTTGTGAGGGTAACAAC  
GGCGGTATGGATGCGGCGGGACCAGAGAAAAATCACTCAGGGTCAATGCCAGCGCTTCGTTAATAC  
AGATGTAGGTGTTCCACAGGGTAGCCAGCAGCATCCTGCGATGCAGATCCGGAACATAATGGTGCA  
GGGCGCTGACTTCCGCGTTTCCAGACTTTACGAAACACGGAACCCGAAGACCATTTCATGTTGTTGC  
TCAGGTGCGCAGACGTTTTGTCAGCAGCAGTCGCTTACGTTTCGCTCGCGTATCGGTGATTCTCTG  
CTAACCAGTAAGGCAACCCCGCCAGCCTAGCCGGGTCTCAACGACAGGAGCAGATCATGCGCAC  
CCGTGGGGCCCGCATGCCGGCGATAATGGCCTGCTTCTCGCCGAAACGTTTGGTGGCGGGACCAGT  
GACGAAGGCTTGAGCGAGGGCGTGCAAGATTCCGAATACCGCAAGCGACAGGCCGATCATCGTCGC  
GCTCCAGCGAAAGCGGTCTCGCCGAAATGACCCAGAGCGCTGCCGGCACCTGTCTACGAGTTG  
CATGATAAAGAAGACAGTCATAAGTGCGGCGACGATAGTCATGCCCCGCGCCACCGGAAGGAGCT  
GACTGGGTGTAAGGCTCTCAAGGGCATCGGTGAGATCCCGGTGCCTAATGAGTGAGCTAACTTAC  
ATTAATTGCGTTGCGCTCACTGCCCGCTTTCAGTCGGGAAACCTGTCTGTCAGCTGCATTAATG  
AATCGGCCAACGCGCGGGGAGAGGCGGTTTGCCTATTGGGCGCCAGGGTGGTTTTTCTTTTACCA  
GTGAGACGGGCAACAGCTGATTGCCCTTACCGCCTGGCCCTGAGAGAGTTGCAGCAAGCGGTCCA  
CGCTGGTTTGGCCCCAGCAGGCGAAAATCCTGTTTGTGTTGGTTAACGGCGGGATATAACATGAGC  
TGTCTTCGGTATCGTCGTATCCCACTACCGAGATATCCGCACCAACGCGCAGCCCGGACTCGGTAA  
TGGCGCGCATTTGCGCCCAGCGCCATCTGATCGTTGGCAACCAGCATCGCAGTGGGAACGATGCCCT  
CATTCAGCATTTGCATGGTTTGTGAAAACCGGACATGGCACTCCAGTCGCCTTCCCGTTCCGCTA  
TCGGCTGAATTTGATTGCGAGTGAGATATTTATGCCAGCCAGCCAGACGCGAGCCGCGGAGACAG  
AACTTAATGGGCCCCGCTAACAGCGCGATTTGCTGGTGACCCAATGCGACCAGATGCTCCACGCCCA  
GTCGCGTACCGTCTTCATGGGAGAAAATAATACTGTTGATGGGTGTCTGGTCAGAGACATCAAGAA  
ATAACGCCGGAACATTAGTGACAGGAGCTTCCACAGCAATGGCATCCTGGTCATCCAGCGGATAGT  
TAATGATCAGCCCACTGACGCGTTGCGCGAGAAGATTGTGCACCGCCGCTTTACAGGCTTCGACGC  
CGTTTCGTTCTACCATCGACACCACCGCTGGCACCCAGTTGATCGGCGCGAGATTTAATCGCCG  
CGACAATTTGCGACGGCGCGTGACGGGCCAGACTGGAGGTGGCAACGCCAATCAGCAACGACTGTT  
TGCCCGCCAGTTGTTGTGCCACGCGGTTGGGAATGTAATTCAGCTCCGCCATCGCCGCTTCCACTT  
TTTCCCGCGTTTTTCGAGAAACGTGGCTGGCCTGGTTTACCACGCGGGAAACGGTCTGATAAGAGA  
CACCGGCATACTCTGCGACATCGTATAACGTTACTGGTTTACATTACCAACCTGAATTGACTCT  
CTTCCGGGCGCTATCATGCCATACCGCGAAAGGTTTTGCGCCATTTCGATGGTGTCCGGGATCTCGA  
CGCTCTCCCTTATGCGACTCCTGCATTAGGAAGCAGCCAGTAGTAGGTTGAGGCCGTTGAGCACC  
GCCGCCGCAAGGAATGGTGCATGCAAGGAGATGGCGCCCAACAGTCCCCCGCCACGGGGCCTGCC  
ACCATACCACGCGGAAACAAGCGCTCATGAGCCCGAAGTGGCGAGCCCGATCTTCCCATCGGTG  
ATGTCGGCGATATAGGCGCCAGCAACCGCACCTGTGGCGCCGGTGATGCCGGCCACGATGCGTCCG

09688672-101000

Figure 2: Nucleotide sequence of TbF15  
Sheet 3 of 4

GCGTAGAGGATCGAGATCTCGATCCCGCGAAATTAATACGACTCACTATAGGGGAATTGTGAGCGG  
ATAACAATTCCCCTCTAGAAATAATTTTGTTTAACTTTAAGAAGGAGATATACATATGGGCCATCA  
TCATCATCATCACGTGATCGACATCATCGGGACCAGCCCCACATCCTGGGAACAGGCGGCGGCGGA  
GGCGGTCCAGCGGGCGCGGGATAGCGTCGATGACATCCGCGTCGCTCGGGTCATTGAGCAGGACAT  
GGCCGTGGACAGCGCCGGCAAGATCACCTACCGCATCAAGCTCGAAGTGTCGTTCAAGATGAGGCC  
GGCGCAACCGAGGTGTGGCTCGAAACCACCGAGCGGTTCCGCTGAAACGGGCGCCGGCGCCGGTAC  
TGTCGCGACTACCCCCGCGTCGTCGCCGGTGACGTTGGCGGAGACCGGTAGCACGCTGCTCTACCC  
GCTGTTCAACCTGTGGGGTCCGGCCTTTACGAGAGGTATCCGAACGTCACGATCACCGCTCAGGG  
CACCGGTTCTGGTGCCGGGATCGCGCAGGCCGCCGGCGGGACGGTCAACATTGGGGCCTCCGACGC  
CTATCTGTGGAAGGTGATATGGCCGCGCACAAGGGGCTGATGAACATCGCGCTAGCCATCTCCGC  
TCAGCAGGTCAACTACAACCTGCCCCGAGTGAGCGAGCACCTCAAGCTGAACGGAAAAGTCTGGC  
GGCCATGTACCAGGGCACCATCAAAACCTGGGACGACCCGCAGATCGCTGCGCTCAACCCCGGCGT  
GAACCTGCCCCGGCACCGCGGTAGTTCCGCTGCACCGCTCCGACGGGTCCGGTGACACCTTCTTGTT  
CACCCAGTACCTGTCCAAGCAAGATCCCGAGGGCTGGGGCAAGTCGCCCCGCTTCGGCACCCCGT  
CGACTTCCCGGCGGTGCCGGGTGCGCTGGGTGAGAACGGCAACGGCGGCATGGTGACCGGTTGCGC  
CGAGACACCGGGCTGCGTGCCCTATATCGGCATCAGCTTCTCGACACCGCCAGTCAACGGGGACT  
CGGCGAGGCCCAACTAGGCAATAGCTCTGGCAATTTCTTGTTGCCCGACGCGCAAAGCATTACGGC  
CGCGGCGGCTGGCTTCGCATCGAAAACCCCGCGCAACCGGCGATTTTCGATGATCGACGGGCCCCG  
CCCGGACGGCTACCCGATCATCAACTACGAGTACGCCATCGTCAACAACCGGCAAAAGGACGCCGC  
CACCGCGCAGACCTTGCAGGCATTTCTGCACTGGGCGATCACCGACGGCAACAAGGCCTCGTTCTT  
CGACCAGGTTTATTTCCAGCCGCTGCCGCCCGCGGTGGTGAAGTTGTCTGACCGGTTGATCGCGAC  
GATTTCCAGCGCTGAGATGAAGACCGATGCCGCTACCTTCGCGCAGGAGGCAGGTAATTTTCGAGCG  
GATCTCCGGCGACCTGAAAACCCAGATCGACCAGGTGGAGTCGACGGCAGGTTTCGTTGCAGGGCCA  
GTGGCGCGGCGCGCGGGGACGGCCGCCAGGCCGCGGTGGTGCGCTTCCAAGAAGCAGCCAATAA  
GCAGAAGCAGGAATCGACGAGATCTCGACGAATATTCGTCAGGCCGGCGTCCAATACTCGAGGGC  
CGACGAGGAGCAGCAGCAGGCGCTGTCTCGCAAATGGGCTTTACTCAGTCGCAGACCGTGACGGT  
GGATCAGCAAGAGATTTTGAACAGGGCCAACGAGGTGGAGGCCCGATGGCGGACCCACCGACTGA  
TGTCCCCATCACACCGTGCGAACTCACGGCGGCTAAAAACGCCGCCCAACAGCTGGTATTGTCCGC  
CGACAACATGCGGGAATACCTGGCGGCCGGTGCCAAAGAGCGGCAGCGTCTGGCGACCTCGCTGCG  
CAACGCGGCCAAGGCGTATGGCGAGGTTGATGAGGAGGCTGCGACCGCGCTGGACAACGACGGCGA  
AGGAACTGTGCAGGCAGAATCGGCCGGGGCGTCGGAGGGGACAGTTCGGCCGAACTAACCGATAC  
GCCGAGGGTGGCCACGGCCGGTGAACCCAACTTCATGGATCTCAAAGAAGCGGCAAGGAAGCTCGA  
AACGGGCGACCAAGGCGCATCGCTCGCGCACTTTGCGGATGGGTGGAACACTTTCAACCTGACGCT  
GCAAGGCGACGTCAAGCGGTTCCGGGGGTTTGACAACCTGGGAAGGCGATGCGGCTACCGCTTGCGA  
GGCTTCGCTCGATCAACAACGGCAATGGATACTCCACATGGCCAAATTGAGCGCTGCGATGGCCAA  
GCAGGCTCAATATGTGCGCGCAGCTGCACGTGTGGGCTAGGCGGGAACATCCGACTTATGAAGACAT  
AGTCGGGCTCGAACGGCTTTACGCGGAAAACCTTCGGCCCGCGACCAAATTCTCCCGGTGTACGC  
GGAGTATCAGCAGAGGTCGGAGAAGGTGCTGACCGAATAACAACAAGGCAGCCCTGGAACCGGT  
AAACCCGCCGAAGCCTCCCCCGCCATCAAGATCGACCCGCCCGCCCTCCGCAAGAGCAGGGATT  
GATCCCTGGCTTCTGATGCCGCCGTCTGACGGCTCCGGTGTGACTCCCGGTACCGGGATGCCAGC  
CGCACCGATGGTTCCGCTACCGGATCGCCGGGTGGTGGCCTCCCGGCTGACACGGCGGCGCAGCT  
GACGTGCGCTGGGCGGGAAGCCGCAGCGCTGTGCGGCGACGTGGCGGTCAAAGCGGCATCGCTCGG  
TGGCGGTGGAGGCGGCGGGGTGCCGTCGGCGCCGTTGGGATCCGCGATCGGGGGCGCCGAATCGGT

09688672-1717000

Figure 2: Nucleotide sequence of TbF15  
Sheet 4 of 4

GCGGCCCCGCTGGCGCTGGTGACATTGCCGGCTTAGGCCAGGGAAGGGCCGGCGGGCGGCCGCGCT  
GGGCGGCGGTGGCATGGGAATGCCGATGGGTGCCGCGCATCAGGGACAAGGGGGCGCCAAGTCCAA  
GGGTTCTCAGCAGGAAGACGAGGCGCTCTACACCGAGGATCGGGCATGGACCGAGGCCGTCATTGG  
TAACCGTCGGCGCCAGGACAGTAAGGAGTCGAAGTGAATTCTGCAGATATCCATCACACTGGCGGC  
CGCTCGAGCACCACCACCACCACCTGAGATCCGGCTGCTAACAAAGCCCGAAAGGAAGCTGAGT  
TGGCTGCTGCCACCGCTGAGCAATAACTAGCATAACCCCTTGGGGCCTCTAAACGGGTCTTGAGGG  
GTTTTTTGCTGAAAGGAGGAACCTATATCCGGAT

09588672-101000



Figure 3: Amino Acid Sequence of TbF14

MQHHHHHTDRVSVGNLRIARVLYDFVNNEALPGTDIDPDSFWAGVDKVVADLTPQNQALLNARDE  
LQAQIDKWHRRRVIEPIDMDAYRQFLTEIGYLLPEPDDFTITTSQVDAEITTTAGPQLVVPVLNAR  
FALNAANARWGSLYDALYGTVDVIPETDGAKEGPTYNKVRGDKVIAYARKFLDDSVPLSSGSFGDAT  
GFTVQDQQLVVALPDKSTGLANPGQFAGYTGAAESPTSVLLINHGLHIEILIDPESQVGTDDRAGV  
KDVI LESA ITTIMDFEDSVAAVDAADKVLGYRNWLGLNKGD LAAAVDKDGTAFRLVLRDRNYTAP  
GGGQFTLPGRSLMFVRNVGHLMTND AIVDTDGSEVFEGIMDALFTGLIAIHGLKASDVNGPLINSR  
TGSIIYIVKPKMHGPAEVAFTCELSRVEDVLGLPQNTMKIGIMDEERRTTVNLKACIKAAADRVVF  
INTGFLDRTGDEIHTSMEAGPMVRKGTMKSPWILAYEDHNVDAGLAAGFSGRAQVGKGMWMTTEL  
MADMVETKIAQPRAGASTAWVPSPTAATLHALHYHQVDVAAVQQGLAGKRRATIEQLLTIPLAKEL  
AWAPDEIREEVDNNCQSILGYVVRWVDQGVGCSKVPDIHDVALMEDRATLRISQLLANWLRHGV  
TSADVRASLERMAPLVDRQNAGDVAYRPMAPNFDDSI AFLAAQELI LSGAQQPNGYTEPILHRRRR  
EFKARAAEKPA PSDRAGDDAARVQKYGGSSVADAERIRRV AERIVATKKQGNDVVVVVSAMGDTTD  
DLLDLAQQVCPAPPPRELDMLLTAGERISNALVAMAIESLGAHARSFTGSQAGVITTGTHGNAKII  
DVT PGRLQTAL EGRVVLVAGFQGV SQDTKDVTTLGRGGS DTTAVAMAAALGADVCEIYTDVDGIF  
SADPRIVRNARKLDTVT FEEMLEMAACGAKVLM LRCVEYARRHNI PVHVRSSYS DRPGTVVVGSIK  
DVP MEDPIL TGVAHDRSEAKVTIVGLPDI PGYAAKV FRAVARRRRQHRHGAAERLQGRGRQDRHHL  
HLLPQTSGPPPWK NWTRSETRSASTQLLYDDHIGKVS LI GAGMRSHPGVTATFCEALAAVGVNIEL  
ISTSEDQRSRCCAATPNWTRPWSRCMKRSGSAATRRPRCTRGRDGRWACQ..

000101" 27988960

Figure 4: Amino Acid Sequence of TbF15

MGHHHHHHVIDIIGTSPTSWEQAAAEAVQRARDSVDDIRVARVIEQDMAVDSAGKITRYIKLEVSF  
KMRPAQPRCGSKPPSGSPETGAGAGTVATTPASSPVTLAETGSTLLYPLFNLWGPAFHRYPNVTI  
TAQGTGSGAGIAQAAAGTVNIGASDAYLSEGDMAAHKGLMNIALAISAQQVNYNLPGVSEHLKLNG  
KVLAAAMYQGTIKTWDDPQIAALNPGVNLPGTAVVPLHRSDGSGDTFLFTQYLSKQDPEGWGKSPGF  
GTTVDFFPAVPGALGENGNMGMTGCAETPGCVAYIGISFLDQASQRGLGEAQLGNSSGNFLLPDAQ  
SIQAAAAGFASKTPANQAIS MIDGPAPDGYPIINYEYAI VNNRQKDAATAQTLQAFLHWAITDGNK  
ASFLDQVHFQPLPPAVVKLS DALIATISSAEMKTDAA TLAEAGNFERISGDLKTQIDQVESTAGS  
LQGQWRGAAGTAAQA AVVRFQEAANKQKQELDEISTNIRQAGVQYSRADEEQQALSSQMGFTQSQ  
TVTVDQQEILNRANEVEAPMADPPTDVPITPCELTA AKNAAQQLVLSADNMREYLAAGAKERQRLA  
TSLRNAAKAYGEVDEEAATALDNDGEGTVQAESAGAVGGDSSAELTDTPRVATAGEPNFMDLKEAA  
RKLETGDQGASLAHFADGWNTFNLT LQGDVKRFRGFDNWE GDAATA CEASLDQQRQWILHMAKLSA  
AMAKQAQYVAQLHVWARREHPTYEDI VGLERLYAENPSARDQILPVYAEYQQRSEKVLTEYNNKAA  
LEPVNPPKPPPAIKIDPPPPPPQE QGLIPGFLMPPSDGSGVT PGTGMPAAPMPVPPTGSPGGGLPADT  
AAQLTSAGREAAALSGDVAVKAASLGGGGGGGVPSAPL GSAIGGAESVRPAGAGDIAGLGQGRAGG  
GAALGGGGMGMPMGAAHQGGGAKSKGSQQEDEALYTEDRAWTEAVIGNRRRQDSKESK.

09588672-101000

Figure 5

	Status	TbF15	TbF6
5004	TB	0.926	1.045
7004	TB	0.928	1.184
9004	TB	1.102	1.365
11004	TB	0.856	1.629
15004	TB	2.035	2.099
17004	TB	2.893	2.867
18004	TB	0.477	0.414
21004	TB	1.062	1.635
23004	TB	0.429	0.501
26004	TB	0.299	0.392
27004	TB	0.244	0.207
28004	TB	2.236	2.04
30004	TB	2.058	1.508
32004	TB	2.324	1.927
33004	TB	1.600	1.578
34004	TB	1.059	1.136
36004	TB	0.546	1.105
37004	TB	1.446	1.989
39004	TB	2.021	2.782
41004	TB	0.511	0.652
43004	TB	0.855	0.483
44004	TB	0.731	0.66
53004	TB	1.100	0.317
FD8-24	Control	0.183	0.314
FD8-25	Control	0.061	0.063
FD8-26	Control	0.066	0.142
FD8-27	Control	0.021	0.115
FD8-28	Control	0.053	0.289
FD8-29	Control	0.114	0.238
FD8-30	Control	0.105	0.146
FD8-31	Control	0.101	0.237
FD8-33	Control	0.080	0.071
FD8-34	Control	0.140	0.117
FD8-35	Control	0.088	0.072
FD8-36	Control	0.081	0.089
FD8-37	Control	0.057	0.06
FD8-38	Control	0.104	0.111
FD8-39	Control	0.221	0.241
FD8-40	Control	0.257	0.265
FD8-41	Control	0.056	0.093
FD8-42	Control	0.184	0.273
FD8-43	Control	0.126	0.126
FD8-44	Control	0.193	0.092
FD8-45	Control	0.058	0.057
FD8-46	Control	0.183	0.23
FD8-48	Control	0.062	0.085
FD8-49	Control	0.134	0.247
Mean		0.113	0.157
SD		0.061	0.086
Mean +3SD		0.298	0.414
Sensitivity		22/23	20/23

CAGGCATGAGCAGAGCGTTTCATCATCGATCCAAAGATGAGGCGCTTGACGAGCTTCTGGGAGTGGAAATACCCAAAGGGGGTATCCT  
GTCCGTACTCGTCTCGCAAGTAGTAGCTAGGTTCCTAGTCACGGTAAGTCCGAACATGCTGGAAGACCCCTAACCTTATGGGTGGTTCCTCCCATAGGA 100

## HTCC-1 FL

M S R A F I I D P T I S A I D G L Y D L L G I G I P N Q G G I L  
TTACTCCTCACTAGAGTACTTCGAAAAAGCCCTGGAGGAGCTGGCAGCAGCGTTTCGGGTGATGGCTGGTTAGGTCGGCCGGGACAAATACGCCGGC  
AATGAGGAGTGATCTCATGAAGCTTTTTCGGGACCTCTCGACCGTCTGCGAAAGGCCACTACCGACCAATCCAAAGCCGGCGCTGTTTATGCGGGCG 200

## HTCC-1 FL

Y S S L E Y F E K A L E E L A A A F P G D G W L G S A A D K Y A G  
AAAAACCGCAACCCAGTGAATTTTTTCCAGGAAGTGGCAGACCTCGATCGTCAGCTCATCAGCCTGATCCACGACCAGGCCAACCGGTCCAGACGACCC  
TTTTTGGCGTTGGTGCACTTAAAAAGGTCTTTGACCGTCTGGAGCTAGCAGTCGAGTAGTCGGACTAGGTGCTGGTCCGGTTGCGCCAGGTCTGCTGGG 300

## HTCC-1 FL

X N R N H V N F F Q E L A D L D R D L I S L I H D Q A N A V Q T T  
GCGACATCCTGGAGGGCGCCAAGAAAGGTCTCGAGTTCGTGCGCCCGTGGCTGTGGACCTGACCTACATCCCGGTCTGTCGGGCACGCCCTATCGGCCGC  
CGCTGTAGGACCTCCCGCGGTTCTTTCCAGAGCTCAAGCACGCGGGCCACCGACACCTGGACTGGATGTAGGGCCAGCAGCCCGTGGGGATAGCCGGCG 400

## HTCC-1 FL

R D I L E G A K K G L E F V R P V A V D L T Y I P V V G H A L S A A  
CTTCCAGGCGCGGTTTTTGGCGGGCGCGATGGCCGTAGTGGGCGGCGCGCTTGCCCTACTTGGTCTGTAAGAACGCTGATCAACCGGACTCAACTCCTCAAA  
GAAGGTCCGCGGCAAAACGCGCCCGCGCTACCGGCATCACCGCGCGCGGAACGGATGAACCAGCACTTTTGGCACTAGTTGCGCTGAGTTGAGGAGTTT 500

## HTCC-1 FL

F Q A P F C A G A M A V V G G A L A Y L V V K T L I N A T Q L L K  
TTGCTTGCCAAATTGGCGGAGTTGGTTCGCGGCCGCCATTGCGGACATCATTTCCGATGTGGCGGACATCATCAAGGGCACCTCGGAGAAGTGTGGGAGT  
AACGAACGGTTTAAACCGCTCAACCAGCGCGGCGGTAACGCCCTGTAGTAAAGCCTACACCGCCTGTAGTAGTTCCCGTGGGAGCCTCTTCACACCTCA 600

## HTCC-1 FL

L L A K L A E L V A A A I A D I I S O V A D I I K G T L G E V W E  
TCATCACAACCGGCTCAACGGCTGAAAGAGCTTTGGGACAAGCTCACGGGGTGGGTGACCGGACTGTTCTCTCGAGGGTGGTCAACCTGGAGTCTCT  
AGTAGTGTTCGCGGAGTTGCCGGACTTTCTCGAAACCTGTTTCGAGTGCCCCACCCACTGGCCTGACAAGAGAGCTCCACCCAGCTTGGACCTCAGGAA 700

## HTCC-1 FL

F I T N A L N G L K E L W D K L T G W V T G L F S R G W S N L E S F  
CTTTGCGGGCGTCCCGGCTTGACCGGCGGACCGCGGCTTGTGCAAGTGAAGTGGCTTGTTCGGTGGCGCCGGTCTGTCCGCATCGTCGGGCTTGGCT  
GAAACGCCCGCAGGGGCGGAAGTGGCCGCGCTGGTTCGCCAAGCAGCGTTCACTGACCGAACAAGCCACGCGCGCCAGACAGCGGTAGCAGCCCGAACC GA 800

## HTCC-1 FL

F A G V P G L T G A T S G L S Q V T G L F G A A G L S A S S G L A

Fig. 6

sheet 1 of 2

CACGCGGATAGCCTGGCGAGCTCAGCCAGCTTCTCCCTGGCGGGCATTGGGGGGGGTCCGGTCTTGGGGGCGCGAGCCTGGCTCAGGTCCATG  
GTGCGCCTATCGGACCGCTCGAGTCGGTGGAAACGGGGGGACCGGCGTAACCCCCCGCCAGGCCAAAAACCCCGAAGCGGCTCGGACCGAGTCCAGGTAC

900

HTCC-1 FL

H A D S L A S S A S L P A L A G I G G G S G F G G L P S L A Q V H

CCGCTCAACTCGGCAGGCGCTACGACCCCGAGCTGATGGCCGGTCCGGCGCGCTGCCGAGCAGGTCCGCGGGCAGTCCGAGCTGGTCTCCGCGCAGGG  
GGCGGAGTTGAGCCGTCCGCGATGCCGGGGCTCGACTACCGGGCCAGCGCGCGGACGGCTCGTCCAGCCGCCCGTCAGCGTCGACCCAGAGCGCGTCCC

1000

HTCC-1 FL

A A S T R Q A L R P R A Q G P V G A A A E Q V G G Q S Q L V S A Q G

TTCCCAAGGTATGGGCGGACCCGTAGGCATGGGCGGCATGCACCCCTCTTCGGGGGGCTCGAAAGGGACGACGACGAAGAAGTACTCGGAAGGCGCGCG  
AAGGGTTCCATACCCCGCTGGGCATCCGTACCCGCGGTACGTGGGGAGAAGCCCCCGCAGCTTTCCTGCTGCTGCTTCTTCATGAGCCTTCCGCGCCGC

1100

HTCC-1 FL

S Q G M G G P V G M G G M H P S S G A S K G T T T K K Y S E G A A

GCGGGCACTGAAGACGCCGAGCGCGCGCCAGTCGAAGCTGACGCGGGCGGTGGGCAAAAGGTGCTGGTACGAAACGTCGCTAACGGCATGGCGAGCCAA  
CGCCCGTGACTTCTGCGGCTCGCGCGCGGTACGCTTCGACTGCGCCCGCCACCCGTTTTCACGACCATGCTTTGCAGCAGATTGCCGTACCGCTCGGTT

1200

HTCC-1 FL

A G T E D A E R A P V E A D A G G G Q K V L V R N V V

0968672-111077

FIG. 6

Sheet 2 of 2

ATGCATCAGCATCACCATCACA TGAGCAGAGCGTTCATCATGATCCAAACGATCAGTGCCATTGACGGCTTGTACGACCTTCTGGGGATTGGAATACCCA  
TACGTAGTGGTAGTGGTAGTGTACTGCTCTGCAAGTAGTAGCTAGGTTGCTAGTCAAGGTAACGCGCAACATGCTGGAAGACCCCTAACCTTATGGGT 100  
M H H H H H H M S R A F I I O P T I S A I O G L Y D L L G I G I P  
ACCAAGGGGTATCCTTTACTCCTCACTAGAGTACTTGGAAAAGCCCTGGAGGAGCTGGCAGCAGCGTTTCCGGGTGATGCGTGGTTAGGTTCCGGCCGC  
TGGTTCCCCCATAGGAAATGAGGAGTATCTCATGAAGCTTTTTCGGGACCTCCTCGACCGTGTGCGCAAGGGCCACTACCGACCAATCCAAGCCGGCG 200  
N O G G I L Y S S L E Y F E K A L E E L A A A F P G D G W L G S A A  
GGACAAATACGCCGGCAAAAACCGCAACGATGAATTTTTTCCAGGAACGGCAGACCTCGATCGTCAGCTCATCAGCGTGATCCAGGACAGGCCAAC  
CCTGTTTATGCGGCCGTTTTTGGCGTTGGTGCACTTAAAAAAGGTCTTGACCGTCTGAGGCTAGCAGTCCAGTAGTGGACTAGGTGCTGGTCCGGTTG 300  
O K Y A G K N R N H V N F F Q E L A O L O R Q L I S L I H O O A N  
GCGGTCCAGACGACCCGCGACATCCTGGAGGGCGCAAGAAGGTCTCGAGTTCTGCGCCCGGTGGCTGTGGACCTGACCTACATCCCGGTGCTCGGGC  
CGCCAGGTCTGCTGGGCGCTGTAGGACCTCCCGCGGTTCTTTCCAGAGCTCAAGCACGCGGGCCACCGACACCTGGACTGGATGTAGGGCCAGCAGCCCG 400  
A V Q T T R O I L E G A K K G L E F V R P V A V D L T Y I P V V G  
ACGCCCTATCGGCCGCTTCCAGGCGCGTTTTGCGGGGCGCGATGGCCGTAGTGGGCGGCGCGCTTGCCCTACTTGGTGGTGA AAAACGCTGATCAACGC  
TGCGGGATAGCCGGCGGAAGGTCCGCGGCAAAAACGCGCCCGCGCTACCGGCATCACCCGCGCGCGGAACGGATGAACAGCACTTTTGGGACTAGTTGCG 500  
H A L S A A F Q A P F C A G A M A V V G G A L A Y L V V K T L I N A  
GACTCAACTCCTCAAATTGCTTGCCAAATTGGCGGAGTTGGTTCGCGGCCGCCATTGCGGACATCATTTCCGATGTGGCGGACATCATCAAGGGCATCCTC  
CTGAGTTGAGGAGTTTAACGAACGGTTTAAACCGCTCAACCGAGCGCGCGGTAACGCGCTGTAGTAAAGCCTACACCGCTGTAGTAGTTCCCGTAGGAG 600  
T Q L L K L L A K L A E L V A A A I A D I I S O V A O L I K G I L  
GGAGAAGTGTGGGAGTTCATCACAACGCGCTCAACGGCTGAAAGAGCTTTGGGACAAGCTCACGGGTGGGTGACCGGACTGTTCTCTCGAGGGTGGT  
CCTCTTCACACCCCTCAAGTAGTGTGTCGCGAGTTGCGGAGCTTTCTCGAAACCTGTTTCGAGTGCCCCACCCACTGGCCTGACAAGAGAGCTCCACCA 700  
G E Y W E F I T N A L N G L K E L W O K L T G W V T G L F S R G W  
CGAACCTGGAGTCTTCTAAGAATTC 726  
GCTTGGACCTCAGGAAGATTCTTAAG  
S N L E S F E F

Fig. 7a

Monday, July 26, 1999 10:50 AM

HTCC-1 (134-37x)

Page 1

HTCC1(134-392) Map.mpd (1 > 361) Site and S

Enzymes : 212 of 515 enzymes (Filtered)

Settings: Linear, Certain Sites Standard Genetic Code

ATGCATCACCATCACCATCAGATGTGGCGGACATCATCAAGGGCATCTCTGGGAGAAGTGTGGGAGTTTCATCACAACGGGCTCAACGGGCTGAAAGAGC  
TACGTAGTGGTAGTGGTAGTGCTACACGGCTGTAGTAGTTCCGGTAGGAGCCTCTTCACACCTCAAGTAGTGTTCGGCGAGTTGCCGGACTTTCTCG  
M H H H H H H D V A O I I K G I L G E V W E F I T N A L N G L K E  
TTTGGGACAAGCTCACCAGGGTGGGTGACCGGACTGTTCTCTCGAGGGTGGTGAACCTGGAGTCTCTTTTGGCGGGTCCCGGGCTTGACCGGGCGGAC  
AAACCTGTTCGAGTCCCGCCACCCACTGGGCTGACAAGAGAGCTCCACCCAGCTTGGACCTCAGGAAGAAAGCGCGGCGAGGGCGGAAGTGGCGCGCTG  
L W D K L T G W V T G L F S R G W S N L E S F F A G V P G L T G A T  
CAGCGGCTTGTGCAAGTGACTGGCTTGTTCGGTGGCGCGGCTGTGTCCGCATCGTGGGCTTGGCTCAGCGGATAGCCTGGCGAGTCAAGCCAGCTTG  
GTGCGCGAAGCAGCGTTCACTGACCGAACAAGCCACGCGCGCCAGACAGGCGTAGCAGCGCGAACCAGAGTGCCTCTCGGACCGCTCGAGTCCGGTGAAC  
S G L S G V T G L F G A A G L S A S S G L A H A D S L A S S A S L  
CCCGCCCTGGCGGGCATTGGGGGCGGGTCCGGTTTTGGGGGCTTGGCGAGCCTGGCTCAGGTCCATGCCCGCTCAACTCGGCAGGCGCTACGGCCCCGAG  
GGGCGGGACCGCGGTAACCCCCGCGCCAGGCCAAAACCCCGGAACGGCTCGGACCGAGTCCAGGTACGGCGGAGTTGAGCGTCCGCGATGCCGGGGCTC  
P A L A G I G G G S G F G G L P S L A Q V H A A S T R Q A L R P R  
CTGATGGCCCCGGTCCGGCGCGCTGCCGAGCAGTCCGGCGGCGAGTCCGAGCTGGTCTCCGCGCAGGGTTCCCAAGGTATGGGCGGACCCGTAGGCATGGG  
GACTACCGGGCCAGCCGCGCGGCGACGGCTCGTCCAGCCGCGCGTCAGCGTCGACCAGAGGCGCGTCCCAAGGGTTCCATACCCGCTGGGCATCCGTACCC  
A O G P V G A A A E Q V G G Q S Q L V S A Q G S Q G M G G P V G M G  
CGGCATGCACCCCTCTTCGGGGGCGTGAAGGGGACGACGACGAAGAAGTACTCGGAAGGCGCGCGCGGGGCACTGAAGACGCCGAGCGCGCGCCAGTC  
GCCGTACGTGGGGAGAAGCCCCCGCAGCTTTCCCTGTGCTGCTTCTTCATGAGCCTTCCGCGCGCGCGCGCGGCTGACTTCTGCGGCTCGCGCGCGGTCAG  
G M H P S S G A S K G T T T K K Y S E G A A A G T E D A E R A P V  
GAAGCTGACGCGGGCGGTGGGCAAAAGGTGCTGGTACGAAACGTCGCTAACGGCGAATTC  
CTTCGACTGCGCGCGCCACCCGTTTTCCACGACCATGCTTTGCAGCAGATTGCCGCTTAAG  
E A D A G G G Q K V L V R N V V R R I

FIG. 7b

Monday, July 26, 1999 10:48 AM  
TCC1(1-129) Map.MPD (1 > 411) Site and sequence  
Enzymes: All 515 enzymes (No  
Settings: Circular, Certain Sites, Standard Genetic Code

H1C-1 (1-129)

Page 1

ATGCATCACCATCACCATCACATGAGCAGAGCGTTCATCATCGATCCAACGATCAGTGCCATTGACGGCTTGACGACCTTCTGGGGATTGG 92  
TACGTAGTGGTAGTGGTAGTGTACTCGTCTCGCAAGTAGTAGCTAGGTGCTAGTCACGGTAACTGCGGAACATGCTGGAAGACCCCTAACC  
M H H H H H H M S R A F I I O P T I S A I O G L Y D L L G I G  
AATACCCAACCAAGGGGGTATCCTTTACTCCTCACTAGAGTACTTCGAAAAAGCCCTGGAGGAGCTGGCAGCAGCGTTTCCGGGTGATGGCT 134  
TTATCGGTTGGTTCCCCCATAGGAAATGAGGAGTGATCTCATGAAGCTTTTTCCGGACCTCCTCGACCGTCTGCGCAAAGGCCCACTACCGA  
I P N Q G G I L Y S S L E Y F E K A L E E L A A A F P G D G  
GGTTAGGTTCCGCCCGCGGACAAATACGCCGGCAAAAAACCGCAACCAACGATGAATTTTTTCCAGGAACCTGGCAGACCTCGATCGTCAGCTCATC 276  
CCAATCCAAGCCGGCGCCTGTTTATGCGGCCGTTTTTGGCGTTGGTGCACCTAAAAAAGGTCTCTGACCGTCTGAGCTAGCAGTCGAGTAG  
W L G S A A O K Y A G K N R N H V N F F Q E L A O L O R Q L I  
AGCCTGATCCACGACCAGGCCAACGCGGTCCAGACGACCCGCGACATCCTGGAGGGCGCCAAGAAAGGTCTCGAGTTCGTGCGCCCGGTGGC 368  
TCGGAAGTAGGTGCTGGTCCGGTTGCGCCAGGTCTGCTGGGCGCTGTAGGACCTCCCGCGGTTCTTTCCAGAGCTCAAGCACGCGGGGCCACCG  
S L I H D G A N A V Q T T R O I L E G A K K G L E F V R P V A  
TGTGGACCTGACCTACATCCCGGTGCTCGGGCACGCCCTATAG 411  
ACACCTGGACTGGATGTAGGGCCAGCAGCCCGTGCGGGATATC  
V O L T Y I P V V G H A L

FIG. 7c

09688672-101000



Monday, July 26, 1999 10:46 AM

2/27/91 TCC1-1022 (1 > 1629) Site and Sequence

Enzymes : All 55 enzymes use the Standard Genetic Code

Linear. Certain Sites Only. Standard Genetic Code

GTATACGTAGTGGTAGTGGTAGTGTGCCGGCCAGGCTATTGAAGGTCGACAGGGTCCCAACCGTGCCTAAGCGGTAAGGCTAGCCCGTCCCGCTACCGCT

Met/HIS TAG T A A S O N F O L S O G G O G F A I P I G O A M A

TCGCGGGCCAGATCCGATCGGGTGGGGGGTCACCCACGGTTCATATCGGGCCFACCGGCTTCCCTGGGCTTGGGTGTTGTGCGACAACAACGGCAACGGCGC  
AGCGCCCGGTCTAGGCTAGCCCAACCCCAAGTGGGTGGCAAGTATAGCCCGGATGGCGGAAGGAGCCGAACCCACAACAGCTGTTGTTGCCGTTGCCGCGC

1 A G Q I R S G G G S P T V H I G P T A F L G L G V V D N N G N G. A

ACGAGTCCAACGGGTGGTCGGGAGCGCTCGGGCGCAAGTCTCGGCATCTCCACCGGCGACGTGATCACC GCGGTCGACGGCGCTCCGATCAACTCGGCC 300  
TGCTCAGGTTGCGCACCCAGCCCTCGCGAGGCGCGCGTTCAGAGCGGTAGAGGTGGCCGCTGCACTAGTGGCGCCAGCTGCCCGAGGCTAGTTGAGCCG

R V Q R V V G S A P A A S L G I S T G D V I T A V O G A P I N S A

ACCGCGATGGCGGACGCGCTTAACGGGCATCATCCCGGTGACGTCATCTCGGTGACCTGGCAAACCAAGTCGGGCGGCACGCGTACAGGGAACGTGACAT  
TGGCGCTACCGCTGCGCGAAATTGCCCGTAGTAGGGCCACTGCAGTAGAGCCACTGGACCGTTTGGTTTCAGCCCGCGTGGCGCATGTCCCTTGCACTGTA

T A M A O A L N G H H P G D V I S V T W Q T K S G G T R T G N V T  
 T G C C C G A G G G A C C C C G G C C G A A T T C C T A G T A C C T A G A G G T T C A A T G A G C A G A G C G T T C A T C G A T C C A A C G A T C A G T G C C A T T G A C G G C T T G T A C G A

TGGCCGAGGGGACCCCCGGCCGAATCTCAAGTACCGTAACTGTTGCTAGTCACGGTAAC TGCCGAACA T GCT  
ACCGGCTCCC TGGGGGCCGGCTTAAGGATCATGGATCTCCAAGTACTCGTCTCGCAAGTAGTAGCTAGGTTCACGGTAAC TGCCGAACA T GCT

Ra12 EcoRI Thrombin hTOC1

C C A G C I D P T I S A I D G L Y O

L A E G P P A E F L V P R G S M S R A F I I D P T I S A I O G L Y U  
CCTTCTGGGGATTGGAATACCCAACCAAGGGGTATCCTTTACTCCTCACTAGAGTACTTCGAAAAAGCCCTGGAGGAGCTGGCAGCAGCGTTTCCGGGT

GGAAGACCCCTAACCTTATGGGTTGGTTCCCCCATAGGAAATGAGGAGTGAATCATGAAGCTTTTGGGACCTCTCTACCCCTGCTGGG  
 hTCC1  
 L L G I G I P N G G I L Y S S L E Y F E K A L E E L A A A F P G

GATGGCTGGTTAGGTTGCGCCGCGGACAAATACGCCGGCAAAAACCGCAACCACGTGAATTTTTTCCAGGAACCTGGCAGACCTCGATCGTCAGCTCATCA  
 CTACCGACCAATCCAAGCGGCGCCTGTTTATGCGGCCGTTTTTGGCGTTGGTGCACCTTAAAAAAGGTCCTTGACCGTCTGGAGCTAGCAGTTCGAGTAGT  
 hTCC1

OGWLGSAADKYAGKNRNHVNF FQEL AOLDRDL  
CGGTCATCCAGGACCAGGCCAACGCGGTCCAGACGACCCGCGACATCCTGGAGGGCGCCAAGAAAGGTCTCGAGTTCGTGCGCCCGGTGGCTGTGGACCT

GCTGATCCACGACCAGGCCAACCGGGTCCAGACGACGCGC  
CGGACTAGGTGCTGGTCCGGTTGCGCCAGGTCTGCTGGGCGCTGAGGACCTCCGCGGTTCTTTCCAGAGCTCAAGCACGCGGGCCACCGACCTGGA

htoc1

N A V G T T R D I L E G A K K G L E F V R P V A V D L

S L I H D Q A N A V G T T R U I L E G A N N G S G  
GACCTACATCCCGGTCGTCGGGCACGCCCTATCGGCCGCCTTCCAGGCGCGCTTTTGCGCGGGCGCGATGGCCGTAGTGGGCGGCGCGCTTGCTACTTG 900  
CTGGATGTAGGGCCAGCAGCCCGTGCGGGATAGCCGGCGGAAGGTCGCGGCAAAACGCGCCCGCGCTACCGGCATCACCCGCCGCGGAACGGATGAAC  
hTCC1

HTCC1  
T Y I P V V G H A L S A A F Q A P F C A G A M A V V G G A L A Y L

FIG. 8

Sheet 1 of 2

1TCGTTGAAAACGCTGATCAACGCGACTCAACTC...AATTGCTTGCCAAATTGGCGGAGTTGGTGGGGGGGGG...GGGACATCATTTGGGATGTGG  
1000  
CAGCACTTTTGGCACTAGTTGGCGTGAGTTGAGGAGTTTAAACGAACGGTTTAAACCGCTCAACGAGCGCGGGCGTAAAGGCTGTAGTAAAGCCTACACC  
hTCC1  
V V K T L I N A T G L L K L L A K L A E L V A A A I A D I I S D V  
CGGACATCATCAAGGGCCTCTCGGAGAAGTGTGGGAGTTTCATCACAACGGCTCAACGGCTGAAAGAGCTTTGGGACAAGCTCACGGGGTGGGTGAC  
1100  
GCCTGTAGTAGTTCCCGTAGGAGCCTCTTACACCCCTCAAGTAGTGTTTGCGCGAGTTGCGCGACTTTCTCGAAACCTGTTCGAGTGCCCCACCCACTG  
hTCC1  
A D I I K G I L G E V W E F I T N A L N G L K E L W D K L T G W V T  
CGGACTGTTCTCTCGAGGGTGGTGAACCTGGAGTCTTCTTTGCGGGCGTCCCCGGCTTGACCGGGCGGACCGAGCGCTTGTGCAAGTGACTGGCTTG  
1200  
GCCTGACAAGAGAGCTCCCACCGAGCTTGGACCTCAGGAAGAAGCGCCCGCAGGGGCGGAACGCGCGCTGGTGGCGGAACAGCGTTCACTGACCGAAC  
hTCC1  
G L F S R G W S N L E S F F A G V P G L T G A T S G L S D V T G L  
TTCGGTGGCGCGGTCTGTCCGCATCGTCGGGCTTGGCTCAGCGGATAGCCTGGCGAGCTCAGCCAGCTTGGCGCCCTGGCGGGCATTGGGGGGGGT  
1300  
AAGCCACGCGCGGCGAGACAGGCGTAGCAGCCGAACCGAGTGGCGCTATCGGACCGCTCGAGTGGTCTGAACGGGCGGGACCGGCGTAACCCCGGCCA  
hTCC1  
F G A A G L S A S S G L A H A D S L A S S A S L P A L A G I G G G  
CCGGTTTTTGGGGGCTTGGCGAGCCTGGCTCAGGTCCATGCCSCCTCAACTCGGCAGGCGCTACGGCCCGAGCTGATGGCCCGGTGGCGCGCGCTGCCGA  
1400  
GGCCAAAACCCCGAACGGCTCGGACCGAGTCCAGGTACGGCGGAGTTGAGCCGTCCCGATGCGGGGGCTCGACTACCGGGCCAGCCGCGCGACGGCT  
hTCC1  
S G F G G L P S L A Q V H A A S T R Q A L R P R A D G P V G A A A E  
GCAGGTGGCGGGCAGTCGCAGCTGGTCTCCGCGCAGGGTTCCCAAGGTATGGGCGGACCGTAGGCATGGGCGGCATGCACCCCTCTTCGGGGGGCTCG  
1500  
CGTCCAGCCGCGCGTCAGCGTCGACCAGAGGCGGTCCCAAGGGTTCATACCCGCTGGGCATCCGTACCCGCGGTACGTGGGGAGAAGCCCCGCGAGC  
hTCC1  
Q V G G Q S Q L V S A Q G S Q G M G G P V G M G G M H P S S G A S  
AAAGGGACGACGACGAAGAAGTACTCGGAAGGCGCGGGCGGGCACTGAAGACGCGGAGCGCGCCAGTCGAAGCTGACGCGGGCGGTGGGCAAAAGG  
1600  
TTTCCCTGCTGCTGCTTCTCATGAGCCTTCCGCGCGCGCGCGCTGACTTCTGCGGCTCGCGCGGCTCAGCTTCGACTGCGCCCGCCACCCGTTTTC  
hTCC1  
K G T T T K K Y S E G A A A G T E O A E R A P V E A D A G G G Q K  
TGCTGGTACGAAACGTCGTCTAAGAATTC 1629  
ACGACCATGCTTTGCAGCAGATTCTTAAG  
hTCC1 EcoRI  
V L V R N V V E F

FIG. B

Sheet 2 of 2

CATATGCATCACCATCACCATCAGATGAGCAGAGCGTTTCATCGATCCAAACGATCAGTGGCATTGACGGCTTGTACGACCTTCGCGGATTGGAATAC  
GTATACGTAGTGGTAGTGGTAGTGTACTCGTCTCGCAAGTAGTAGGTGGTCTAGTACGGTAAGTGGCGAACATGCTGGAAGACCCCTAACCTTATG 100  
Met / HIS TAG HTCC1  
H M H H H H H M S R A F I I O P T I S A I O G L Y O L L G I G I  
CCAACCAAGGGGGTATCCCTTACTCCTCACTAGAGTACTTCGAAAAAGCCCTGGAGGAGCTGGCAGCAGCGTTTCGGGTGATGGCTGGTTAGGTTCCGGC 200  
GGTTGGTTCCCCCATAGGAAATGAGGAGTGATCTCATGAAGCTTTTCGGGACCTCCTCGACCGTCTGTCGCAAGGCCCACTACCGACCAATCCAAGCCG  
HTCC1  
P N O G G I L Y S S L E Y F E K A L E E L A A A F P G O G W L G S A  
CGCGGACAAATACGCCGGCAAAAACCGCAACCACGTGAATTTTTTCAGGAACCTGGCAGACCTCGATCGTCAGCTCATCAGCCTGATCCAGCAGCAGGCC 300  
GCGCCTGTTTATGCGGCGGTTTTGGCGTTGGTGCACCTTAAAAAGGTCTTGACCGTCTGGAGCTAGCAGTCTGAGTCTGGACTAGGTGCTGGTCCGG  
HTCC1  
A O K Y A G K N R N H V N F F O E L A O L O R O L I S L I H O O A  
AACCGCGTCCAGACGACCCGCGACATCCTGGAGGGCGCCAAAGAAAGGTCTCGAGTTCTGCGCCCGGTGGCTGTGGACCTGACCTACATCCCGGTGCTCG 400  
TTGCGCCAGGTCTGCTGGGCGCTGTAGGACCTCCGCGGTTCTTTTCAGAGCTCAAGCAGCGGGGCCACCGACACCTGGACTGGATGTAGGGCCAGCAGC  
HTCC1  
N A V Q T T R O I L E G A K K G L E F V R P V A V O L T Y I P V V  
GGCAGCGCCTATCGGCCGCTTCCAGGCGCGTTTTGCGCGGCGCGATGGCGCTAGTGGGCGGCGCGCTTAAGCTTGCCCTACTTGGTCTGAAAAAGCT 500  
CCGTGCGGGATAGCGGCGGAAGGTCCGCGGCAAAACGCGCGCGCGCTACCGGCATCAGCGCGCGCGAATTGGAACGGATGAACAGCACTTTTGGGA  
HTCC1 Hind3 DELETED  
G H A L S A A F O A P F C A G A M A V V G G A L K L A Y L V V K T L  
GATCAACGCGAAGCTTACTCAACTCCTCAAATTGCTTGCCAAATTGGCGGAGTTGGTGGCGGCGGCCATTGGCGGACATCATTTCGGATGTGGCGGACATC 600  
CTAGTTGCGCTTCGAATGAGTTGAGGAGTTTAACGAACGGTTTAACCGCCTCAACCAGCGCGGCGGTAACGCTGTAGTAAAGCCTACACCGCCTGTAG  
DELETED Hind3 HTCC1  
I N A K L T O L L K L L A X L A E L V A A A I A D I I S O V A O I  
ATCAAGGGCATCCTCGGAGAAGTGTGGGAGTTCATCAGAAACGCGCTCAACGGCCTGAAAGAGCTTTGGGACAAGCTCACGGGGTGGGTGACCGGACTGT 700  
TAGTTCCCGTAGGAGCCTCTTCACACCTCAAGTAGTGTTCGCGGAGTTGCGGACTTTCTCGAAACCTGTTTCGAGTGCCCCACCCACTGGCCTGACA  
HTCC1  
I K G I L G E V W E F I T N A L N G L K E L W O K L T G W V T G L  
TCTCTGAGGGTGGTGAACCTGGAGTCTCTTTTGGGGGCTCCCGGCTTGACCGGCGCGACCAGCGGCTTGTCGCAAGTGAAGTGGCTTGTTCGGTGC 800  
AGAGAGCTCCACCAAGCTTGGACCTCAGGAAGAAACGCGCAGGGGCGCAACTGGCGCGCTGGTTCGCGAAGCAGCGTTCACTGACCGAACAAGCCAGC  
HTCC1  
F S R G W S N L E S F F A G V P G L T G A T S G L S O V T G L F G A  
GGCGGCTGTGTCGCATCGTGGGCTTGGCTCACGCGGATAGCCTGGCGAGCTCAGCCAGCTTGCCCGCCTGGCGGCGATTGGGGGCGGGTCCGGTTTT 900  
CCGGCCAGACAGGCGTAGCAGCCCGAACCAGTGCCTATCGGACCGCTCGAGTCGGTGAACGGGCGGGACCGGCGTAACCCCGCCAGGCCAAAA  
HTCC1  
A G L S A S S G L A H A O S L A S S A S L P A L A G I G G G S G F

FIG. 9a

Sheet 1 of 2

3GGGGCTTGGCGAGCCTGGGTCAGGTCCATGCCCAACTGGGCAGGCGCTACGGCCCCGAGCTGATGGCCGGCGCGCTGCCGAGCAGGTGG 1000  
CCCCCGAACGGCTCGGACCGAGTCCAGGTAAGGGGAGTTGAGCCGTCCGCGATGCCGGGGCTCGACTACCGGGCCAGCCCGCGGACGGCTCGTCCAGC  
HTCC1  
G G L P S L A Q V H A A S T R C A L R P R A D G P V G A A A E Q V  
GCGGGCAGTGGCAGCTGGTCTCCGCGCAGGCTTCCCAAGGTATGGGCGGACCCGTAGGCATGGGCGGCATGCACCCCTCTTCGGGGGCGTCGAAAGGGAC 1100  
CGCCCGTCAGCGTCGACCGAGAGGCGCGTCCCAAGGTTCCATACCCGCTGGGCATCGGTACCCGCGCTACGTGGGGAGAAGCCCCCGCAGCTTTCCCTG  
HTCC1  
G G Q S G L V S A Q G S G G M G G P V G M G G M H P S S G A S K G T  
GACGACGAAGAAGTACTCGGAAGGCGCGGCGGCGGCGCACTGAAGACGCCGAGCGCGCGCAGTCGAAGCTGACGCGGGCGGTGGGCAAAAGGTGCTGGTA 1200  
CTGCTGCTTCTTCATGAGCCTTCCGCGCGCGCGCGGTGACTTCTCGGGCTCGCGCGCGGTGAGCTTCCGACTGCGCCCCGCCACCCGTTTTCCACGACCAT  
HTCC1  
T T K K Y S E G A A A G T E D A E R A P V E A D A G G G Q X V L V  
CGAAACGTCGTCTAACGGCGAATTC 1225  
GCTTTGCAGCAGATTGCCGCTTAAG  
HTCC1 EcoRI  
R N V V . R R I

FIG. 9a

Sheet 2 of 2

09688672-1717000

Friday, February 19, 1999 2:28 PM  
 HCC1 peptides seq Map (1 > 1179) Site and Sequence  
 Enzymes: All 470 enzymes (No Filter)  
 Settings: Linear, Certain Sites Only, Standard Genetic Code

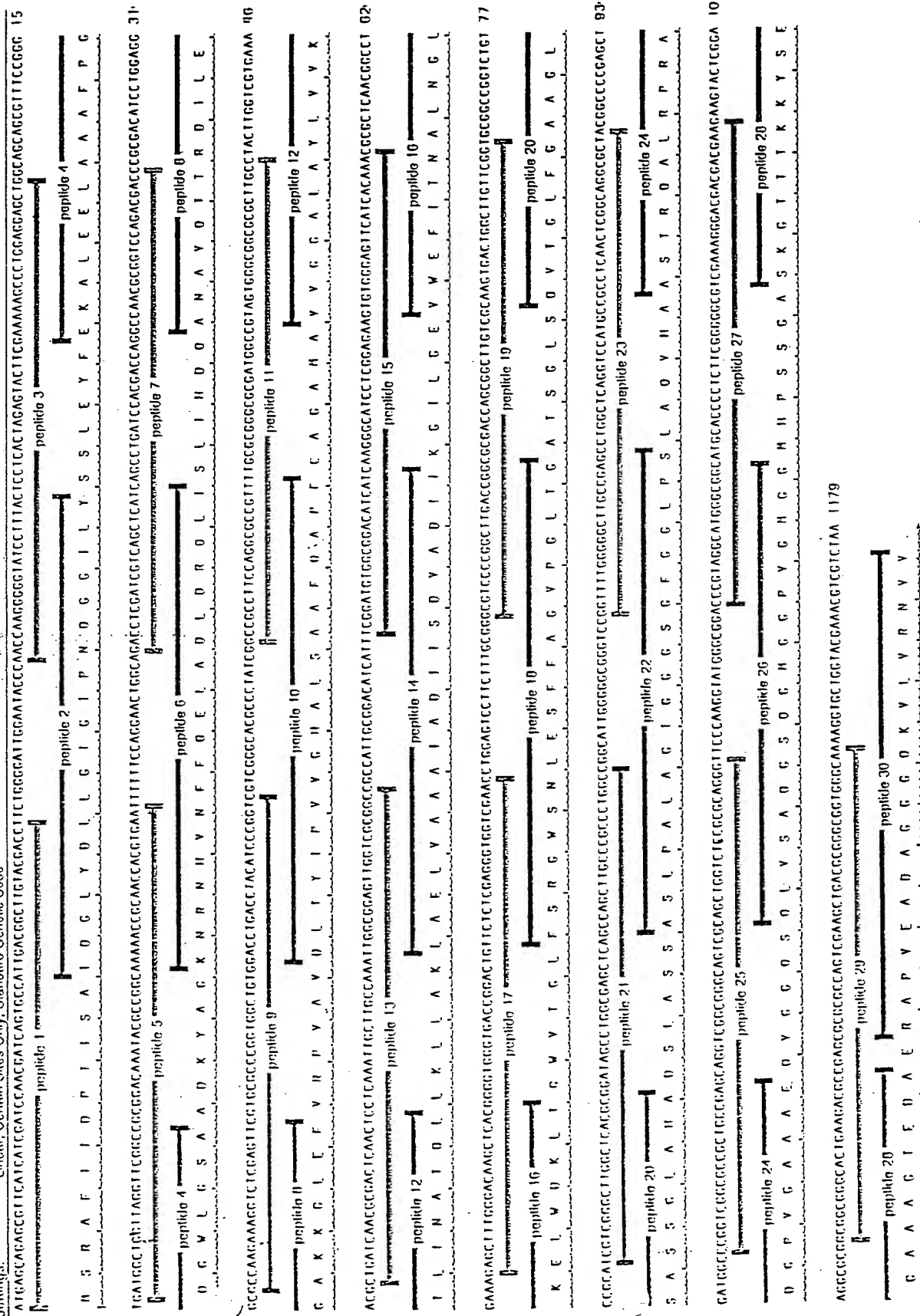


Fig. 96

# T Cell Epitope Mapping of HTCC-1

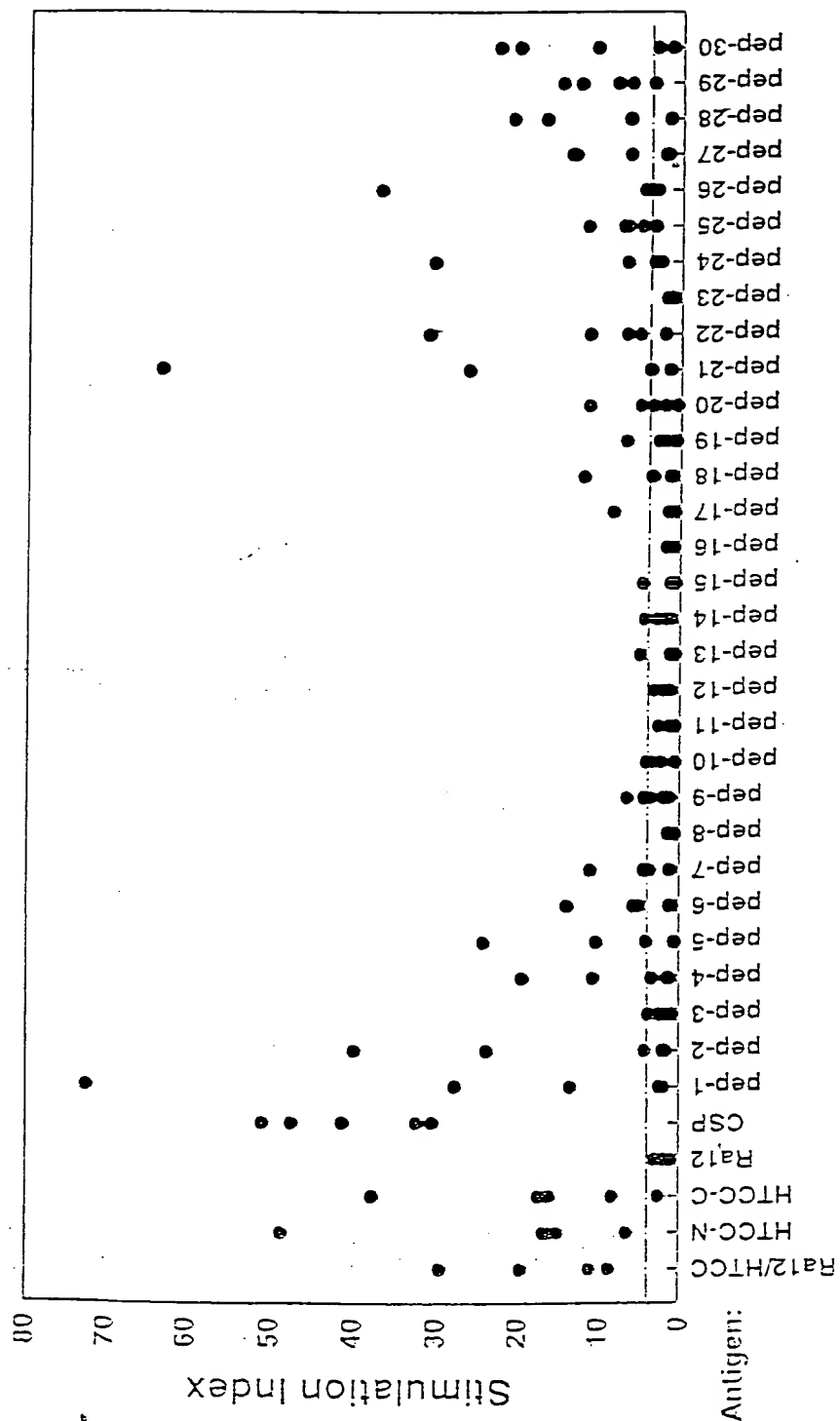


FIG. 9c

CATATGCATCACCATCACCATCACATGAGCAGAGGCTTCATCATCGATCCAAACGATCAGTGGCAATCAGGCTTGACGACCTTCTGGGGA  
GTATACGTAGTGGTAGTGGTAGTGTACTCGTCTCGCAAGTAGTAGCTAGGTTGCTAGTCACGGTAACGCGCAACATGCTGGAAGACCCCT

91

hTCC1

H M H H H H H H M S R A F I I O P T I S A I D G L Y O L L G

TTGGAATACCCAACCAAGGGGGTATCCTTTACTCCTCACTAGAGTACTTCGAAAAAGCCCTGGAGGAGCTGGCAGCAGCGTTTCCGGGTGA  
AACCTTATGGGTTGGTTCCCCCATAGGAAATGAGGAGTGATCTCATGAAGCTTTTTCGGGACCTCCTCGACCGTCTGCGCAAAGGCCCACT

182

hTCC1

I G I P N Q G G I L Y S S L E Y F E K A L E E L A A A F P G O

TGGCTGGTTAGGTTTCGGCCGCGGACAAATACGCCGCAAAAAACCGCAACCACGTGAATTTTTTCCAGGAACCTGGCAGACCTCGATCGTCAG  
ACCGACCAATCCAAGCCGCGCCTGTTTATGCGGCCGTTTTTGGCGTTGGTGCACCTTAAAAAAGGTCTTGACCGTCTGGAGCTAGCAGTC

273

hTCC1

G W L G S A A D K Y A G K N R N H V N F F Q E L A D L O R Q

CTCATCAGCCTGATCCACGACCAGGCCAACGCGTCCAGACGACCCGCGACAAGCTTATCCTGGAGGGGCGCAAGAAAGGTCTCGAGTTCC  
GAGTAGTCGGACTAGGTGCTGGTCCGGTTGCGCCAGGTCTGCTGGGCGCTGTTTGAATAGGACCTCCCGCGTTCTTTCCAGAGCTCAAGC

364

hTCC1

Hind3

DELETED

L I S L I H O Q A N A V Q T T R O K L I L E G A K K G L E F

TGCGCCCGGTGGCTGTGGACCTGACCTACATCCCGTCTGCGGGCAGCCCTATCGCCCGCCTTCCAGGCGCCGTTTTGCGCGGGCGCGAT  
ACGCGGGCCACCGACACCTGGACTGGATGTAGGGCCAGCAGCCCGTGCGGGATAGCCGGCGGAAGGTCCGCGGCAAAACGCGCCCGCGCTA

455

DELETED

V R P V A V O L T Y I P V V G H A L S A A F Q A P F C A G A M

GGCCGTAGTGGGCGCGCGCTTGCTACTTGGTCGTGAAAAACGCTGATCAACGCGACTCAACTCCTCAAATTGCTTGCCAAATTGGCGGAG  
CCGGCATCACCCGCGCGCGGAACGGATGAACCAGCACTTTTGGGACTAGTTGCGCTGAGTTGAGGAGTTTAAACGAACGGTTTAAACGCGCTC

546

DELETED

A V V G G A L A Y L V V K T L I N A T Q L L K L L A K L A E

TTGGTTCGCGGCCGCCATTGCGGACATCATTTCCGATGTGGCGGACATCATCAAGGGCATCCTCGGAGAAGTGTGGGAGTTTCATCACAAACG  
AACCAGCGCCGGCGGTAACGCTGTAGTAAAGCCTACACCGCCTGTAGTAGTTCCCGTAGGAGCCTCTTCACACCCTCAAGTAGTGTTCG

637

DELETED

L V A A A I A D I I S D V A D I I K G I L G E V W E F I T N

CGAAGCTTCTCAACGGCCTGAAAGAGCTTTGGGACAAGCTCACGGGGTGGGTGACCGGACTGTTCTCTCGAGGGTGGTTCGAACCTGGAGTC  
GCTTCGAAGAGTTGCCGGACTTTCTCGAAACCCTGTTTCGAGTGCCCCACCCACTGGCCTGACAAGAGAGCTCCCACCAGCTTGGACCTCAG

728

Hind3

hTCC1

A K L L N G L K E L W D K L T G W V T G L F S R G W S N L E S

CTTCTTTGCGGGCGTCCCGGCTTGACCGGCGGACCGAGCGGCTTGTGCAAGTGACTGGCTTGTTCGGTGCAGCGGGTCTGTCCGCATCG  
GAAGAAACGCCCGCAGGGGCCGAACCTGGCCGCGTGGTTCGCCGAACAGCGTTCACTGACCGAACAAGCCACGCGCGCCAGACAGCGGTAGC

819

hTCC1

F F A G V P G L T G A T S G L S Q V T G L F G A A G L S A S

FIG. 9d

Sheet 1 of 2

TCGGGCTTGCTCAGCGGATAGCCTGG CTGAGCCAGCTTGCCCGCCCTGGCCGGCATTGG GGTCCGGTTTTGGGGGCTTGC  
AGCCCGAACCAGTCCGCCCTATCGGACCGCTCGAGTCGGTCTGAACGGGCGGGACCGGCCSTAACCCCGCCAGGCCAAAACCCCGAACC 910  
hTCC1  
S G L A H A D S L A S S A S L P A L A G I G G G S G F G G L  
CGAGCCTGGCTCAGGTCCATGCCGGCTCAACTCGGCAGGCGCTACGGCCCCGAGCTGATGGCCCGGTGGCGCCGCTGCCGAGCAGGTCCG 1001  
GCTCGGACCGAGTCCAGGTACGGCGGAGTTGAGCCGTCCGCGATGCCGGGGCTCGACTACCGGGCCAGCCGGCGGACGGCTCGTCCAGCC  
hTCC1  
P S L A Q V H A A S T R Q A L R P R A D G P V G A A A E Q V G  
CGGGCAGTCGCAGCTGGTCTCCGCGCAGGGTCCCAAGGTATGGGCGGACCCGTAGGCATGGCCGGCATGCACCCCTCTTCGGGGGCGTCCG 1092  
GCCCGTCAGCGTCGACCAGAGGCGCGTCCCAAGGTTCCATACCCGCTGGGCATCCGTACCCGCGGTACGTGGGGAGAAGCCCCCGCAAC  
hTCC1  
G Q S Q L V S A Q G S G G M G G P V G M G G M H P S S G A S  
AAAGGGACGACGACGAAGAAGTACTCGGAAGGCGCGCGCGGGGCACTGAAGACGCGGAGCGCGCGCCAGTCGAAGCTGACGCGGGCGGTG 1183  
TTTCCCTGCTGCTGCTTCTTCATGAGCCTTCCGCGCGCGCGCCCGTGACTTCTCGGGCTCGCGCGCGGTGAGCTTCGACTGCGCCCGCCAC  
hTCC1  
K G T T T K K Y S E G A A A G T E D A E R A P V E A D A G G  
GGCAAAAGGTGCTGGTACGAAACGTCGTCTAACGGCGAATTC 1225  
CCGTTTTCCACGACCATGCTTTGCAGCAGATTGCCGCTTAAG  
hTCC1 EcoRI  
G Q K V L V R N V V R R I

FIG. 9d  
Sheet 2 of 2



Monday, July 25, 1999 3:25 PM

ht(184-392)-H9-ht(1-129).mpd (1 > 2232) S1 c Sequence

Enzymes : 3 of 515 enzymes (Filtered)

Settings: Linear, Certain Sites Standard Genetic Code

71274134-37X-1911-111-111-111 Page 1

000107-2298960

CATATGCATCACCATCACCATCAGATGTGGCGGACATCATCAAGGGCATCTCTGGAGAAAGTGTGGAGTTTCATCACAACGGCTCAACGGCGTGAAG  
GTATACGTAGTGGTAGTGGTAGTGTCTACACGGCTGTAGTAGTTCCTCGTAGGAGCCTCTTCAACGGCTCAAGTAGTGTTCGGCGAGTTGCCGGACTTTC 100  
Met / HIS TAG hTCC1 (184-392)  
H M H H H H H H D V A D I I K G I L G E V W E F I T N A L N G L K  
AGCTTTGGGACAAGCTCAGCGGGGTGGGTGACGGGACTGTTCTCTCGAGGGTGGTTCGAACCTGGAGTCTTCTTTGGCGGCTGCCCGGCTTGACCGGGCG 200  
TCGAAACCTGTTCGAGTGCCTCCACCCACTGGCTGACAAGAGAGCTCCACAGCTTGGACCTCAGGAAGAAACGCCCGAGGGGCCGAAGTGGCCGCG  
hTCC1 (184-392)  
E L W D K L T G W V T G L F S R G W S N L E S F F A G V P G L T G A  
GACCAGCGGCTTGTGCAAGTGAAGTGGCTTGTTCGGTGGCGCGGTCTGTCCGCATCGTGGGCTTGGCTCAGCGGATAGCCTGGCGAGCTCAGCCAGC 300  
CTGGTCGCCGAACAGCGTTCACTGACCGAACAAGCCAGCGCGCCAGACAGGCGTAGCAGCCGAAACGAGTGGCGCTATCGGACCGCTCGAGTCGGTCG  
hTCC1 (184-392)  
T S G L S O V T G L F G A A G L S A S S G L A H A O S L A S S A S  
TTGCCCGCCCTGGCGGCAATTGGGGGCGGGTCCGGTTTGGGGGCTTCCCGAGCCTGGCTCAGGTCCATGCCGCTCAACTCGGCAGGCGCTACGGCCCC 400  
AACGGGCGGGACCGGCGTAACCCCGCGCCAGGCCAAAACCCCGGAACGGCTCGGACCGAGTCCAGGTACGGCGGAGTTGAGCCGTCCGCGATGCCGGGG  
hTCC1 (184-392)  
L P A L A G I G G G S G F G G L P S L A Q V H A A S T R Q A L R P  
GAGCTGATGGCCCGGTTCGGCGCCGCTGCCGAGCAGGTGGCGGGCAGTGCAGCTGGTCTCCGCGCAGGGTTCCCAAGGTATGGCGGACCCGTAGGCAT 500  
CTCGACTACCGGGCCAGCCGCGGCGACGGCTGTCCAGCGCGCCGTACCGCTCGACCGAGGGCGGTCCCAAGGGTTCCATACCCGCTGGGCATCCGTA  
hTCC1 (184-392)  
R A D G P V G A A A E Q V G G Q S Q L V S A Q G S Q G M G G P V G M  
GGCGGCATGCACCCCTCTTCGGGGGCGTCAAAAGGGACGACGACGAAGAAGTACTCGGAAGGGCGCGGGCGGGCACTGAAGACGCCGAGCGCGGCCA 600  
CCGCGCGTACGTGGGGAGAAGCCCCCGCAGCTTTCCCTGCTGCTGCTTCTTCATGAGCCTTCCGCGCGCGCGCGGTGACTTCTGCGGCTCGCGCGCGGT  
hTCC1 (184-392)  
G G M H P S S G A S K G T T T K K Y S E G A A A G T E D A E R A P  
GTCGAAGCTGACGGGGCGGTGGGCAAAAGGTGCTGGTACGAAACGTCGTCGAATTCATGGTGGATTTGGGGCGGTTACCACCGGAGATCAACTCCGCGA 700  
CAGCTTCGACTGCGCGCGCCACCGTTTTCCACGACCATGCTTTGCAGCAGCTTAAGTACCACCTAAAGCCCCGCAATGGTGGCCTCTAGTTGAGGCGCT  
hTCC1 (184-392) EcoRI TbH9  
V E A D A G G S Q K V L V R N V V E F M V D F G A L P P E I N S A  
GGATGTACGCGGGCGGGTTCCGGCTCGCTGGTGGCGCGGCTCAGATGTGGGACAGCGTGGCGAGTGACCTGTTTTGGCGCGCTCGGCGTTTCAGTC 800  
CCTACATGCGCGCGGCCCAAGCGGAGCGACCAACCGCGCGGAGTCTACACCTGTGCGACCGCTCACTGGACAAAAGCGCGCGAGCCGCAAGTCAAG  
TbH9  
R M Y A G P G S A S L V A A A Q M W D S V A S O L F S A A S A F Q S  
GGTGGTCTGGGGTCTGACGGTGGGGTGGTGGATAGGTTCTGTCGGCGGGTCTGATGGTGGCGGGCGCTCGCGGTATGTGGCGTGGATGAGCGTCACCGCG 900  
CCACCAGACCCAGACTGCCACCCAGCACCTATCCAAGCAGCCGCCAGACTACCACCGCGCGGAGCGGCATACACCGACCTACTCGCAGTGGCGC  
TbH9  
V V W G L T V G S W I G S S A G L M V A A A S P Y V A W M S V T A

FIG. 10

Sheet 1 of 3

Sheet 2 of 3

GGATTGGAATACCCAACCAAGGGGGTATCTT CCTCACTAGAGTACTTCGAAAAAGCCCTGGAGGAGCT GCAGCGTTTCGGGGTGATGGCTG  
CCTAACCTTATGGGTTGGTTCCCCCATAGGAAATGAGGAGTGATCTCATGAAGCTTTTTCGGGACCTCCTCGACCGTCTGCGCAAGGCCCACTACCGAC 2000  
hTCC1 (1-129)  
G I G I P N Q G G I L Y S S L E Y F E K A L E E L A A A F P G D G W  
GTTAGGTTGGGCGCGGACAAATACGCCGGCAAAAACCGCAACCACGTGAATTTTTTCCAGGAACCTGGCAGACCTCGATCGTCAGCTCATCAGCCTGATC 2100  
CAATCCAAGCCGGCGCCTGTTTATGCGGCGGTTTTTGGCGTTGGTGCACCTAAAAAAGGTCTTGACCGTCTGGAGCTAGCAGTCGAGTAGTCGGACTAG  
hTCC1 (1-129)  
L G S A A D K Y A G K N R N H V N F F Q E L A D L O R O L I S L I  
CACGACCAGGCCAACGCCGTCCAGACGACCCGCGACATCCTGGAGGGCGCCAAGAAAGGTCTCGAGTTCTGCGCCCGGTGGCTGTGGACCTGACCTACA 2200  
GTGCTGGTCCGGTTGCGCCAGGTCTGCTGGGCGCTGTAGGACCTCCCGCGGTTCTTTCCAGAGCTCAAGCACGCGGGCCACCGACACCTGGACTGGATGT  
hTCC1 (1-129)  
H O O A N A V O T T R O I L E G A K K G L E F V R P V A V O L T Y  
TCCCGGTCGTCGGGCACGCCCTATAAGATATC 2232  
AGGGCCAGCAGCCCGTGCGGGATATTCTATAG  
hTCC1 (1-129) RV  
I P V V G H A L D I

FIG. 10

Sheet 3 of 3

Monday, July 26, 1999 2:42 PM

ht(1-149)-H9-ht(161-392).mpd (1 > 2365) Site and Sequence

Enzymes : 3 of 515 enzymes (Filtered)

Settings : Circular, Certain S Only, Standard Genetic Code

Page 1

CATATGCATCACCATCACCATCACCATGAGGAGGCTTCATCATCATCAACGATCAGTGCCATGGGCTGTACGACCTTCTGGGGA  
GTATACGTAGTGGTAGTGGTAGTGTACTCGTCTCGCAAGTAGTAGCTAGGTTGCTAGTCACGGTAACTGCCGAACATCTCGGAAGACCCCT 91  
Met / HIS TAG HTCC1 (1-149)  
H M H H H H H H M S R A F I I O P T I S A I D G L Y D L L G  
TTGGAATACCCAACCAAGGGGGTATCCTTTACTCCTCACTAGAGTACTTCGAAAAAGCCCTGGAGGAGCTGGCAGCAGCGTTTCCGGGTGA  
AACCTTATGGGTTGGTTCCCCCATAGGAAATGAGCAGTATCTCATGAAGCTTTTTTCGGGACCTCCTCGACCGTCTGCGCAAAGGCCCACT 132  
HTCC1 (1-149)  
I G I P N Q G G I L Y S S L E Y F E K A L E E L A A A F P G D  
TGGCTGGTTAGGTTTCGGCCGCGGACAAATACGCCGGCAAAAACCGCAACCACGTGAATTTTTTCCAGGAACGGCAGACCTCGATCGTCAG  
ACCGACCAATCCAAGCCGGCGCCTGTTTATGCGGCCGTTTTTGGCGTTGGTGCACTAAAAAAGGTCCTTGACCGTCTGGAGCTAGCAGTC 273  
HTCC1 (1-149)  
G W L G S A A D K Y A G K N R N H V N F F Q E L A D L D R Q  
CTCATCAGCCTGATCCACGACCAGGCCAACGCCGTCCAGACGACCCGCGACATCCTGGAGGGCGCCAAGAAAGGTCTCGAGTTCTGTGCCG  
GAGTAGTCGGAAGTAGGTGCTGGTCCGGTTGCGCCAGGTCTGCTGGGCGCTGTAGGACCTCCCGCGGTTCTTTCCAGAGCTCAAGCAGCGG 364  
HTCC1 (1-149)  
L I S L I H D Q A N A V Q T T R D I L E G A K K G L E F V R  
CGGTGGCTGTGGACCTGACCTACATCCCGGTGCTCGGCACGCCCTATCGGCCGCTTCCAGGCGCGTTTTTGGCGGGCGCGATGGCCGT  
GCCACCGACACCTGGACTGGATGTAGGGCCAGCAGCCCGTGCGGGATAGCCGGCGGAAGGTCCGCGGCAAAACGCGCCCGCGCTACCGGCA 455  
HTCC1 (1-149)  
P V A V D L T Y I P V V G H A L S A A F Q A P F C A G A M A V  
AGTGGGCGGCGCGCTTAAGCTTATGGTGGATTTTCGGGGCGTTACCACCGGAGATCAACTCCGCGAGGATGTACGCCGGCCCGGGTTCCGGC  
TCACCCGCGCGCGAATTGGAATACCACCTAAAGCCCCGCAATGGTGGCCTCTAGTTGAGGCGCTCTACATGCGGCCGGGCCCAAGCCGG 546  
HTCC1 (1-149) Hind3 TbH9  
V G G A L K L M V D F G A L P P E I N S A R M Y A G P G S A  
TCGCTGGTGGCCGCGGCTCAGATGTGGGACAGCGTGGCGAGTGACCTGTTTTTCGGCCGCGTCCGCGTTTTCACTCGGTGGTCTGGGGTCTGA  
AGCGACCACCGCGCGCGAGTCTACACCTGTGCGACCGCTCACTGGACAAAAGCCGGCGCAGCCGCAAGTCAGCCACCAGACCCCACT 637  
TbH9  
S L V A A A O M W D S V A S D L F S A A S A F Q S V V W G L  
CGGTGGGGTCTGTGGATAGGTTCTGTCGGCGGGTCTGATGGTGGCGGCGGTCTCGCCGATGTGGCGTGGATGAGCGTCACCGCGGGGCGAGGC  
GCCACCCAGCACCTATCCAAGCAGCCGCCAGACTACCACCGCCGCCAGAGCGGCATACACCGCACCTACTCGCAGTGGCGCCCGGTCCG 728  
TbH9  
T V G S W I G S S A G L M V A A V S P Y V A W M S V T A G Q A  
CGAGCTGACCGCCGCCCCAGGTCCGGGTTGCTGCGGCGGCTACGAGACGGCGTATGGGCTGACGGTGGCCCCGCGGTGATCGCCGAGAAC 819  
GCTCGACTGGCGGCGGGTCCAGGCCAACGACGCCGCCGATGCTCTGCCGCATACCCGACTGCCACGGGGGCGGCCACTAGCGGCTCTTG  
TbH9  
E L T A A Q V R V A A A A Y E T A Y G L T V P P P V I A E N

FIG. 11

Sheet 1 of 3

CGTGCTGAACATGATGATTCTGATAGCGA...CTCTTGGGGCAAAACACCCCGGCGATCGCGGT...AGGCCGAATACGGCGAGATGT 910  
GCACGACTTGACTACTAAGACTATCGCTGGTTGGAGAACCCCGTTTGTGGGGCCGCTAGCGCCAGTTCTCCGCTTATGCCGCTCTACA  
TbH9  
R A E L M I L I A T N L L G Q N T P A I A V N E A E Y G E M  
GGGCCCAAGACGCCGCCGCGATGTTTGGCTACGCCCGCGCGACGGCGACGGCGACGTTGCTGCGGTTCGAGGAGGCGCGCGGAGAT 1001  
CCCGGGTTCTGCGGGGGCGCTACAAACCGATGCGGCGCGCGTGGCGCTGCCGCTGCCGCTGCAACGACGGCAAGCTCCTCCGCGGCTCTA  
TbH9  
W A Q D A A A M F G Y A A A T A T A T A T L L P F E E A P E M  
GACCAGCGCGGGTGGGCTCCTCGACAGGCCCGCGCGGTGAGGAGGCTCCGACACCGCGCGCGGAACCAAGTTGATGAACAATGTGCC 1092  
CTGGTCGCGCCACCCGAGGAGCTCGTCCGGCGCGCCAGCTCCTCCGAGGCTGTGGCGCGCGCTTGGTCAACTACTTGTACACGGG  
TbH9  
T S A G G L L E Q A A A V E E A S O T A A A N Q L M N N V P  
CAGGCGCTGCAACAGCTGGCCACGCCACGCAGGCGACCCAGCCTTCTTCCAAGCTGGGTGGCTGTGGAAGACGGTCTCGCCGCATCGGT 1183  
GTCCGCGACGTTGTCGACCGGGTCGGGTGCGTCCCGTGGTGGGAAGAAGTTTCGACCCACCGGACACCTTCTGCCAGAGCGGCGTAGCCA  
TbH9  
Q A L Q Q L A Q P T Q G T T P S S K L G G L W K T V S P H R  
CGCCGATCAGCAACATGGTGTGATGGCCAACAACCACATGTGATGACCAACTCGGGTGTGTGATGACCAACACCTTGAGCTCGATGT 1274  
GCGGCTAGTCGTTGTACACAGCTACCGTTGTTGGTGTACAGCTACTGGTTGAGCCACACAGCTACTGGTTGTGGAACCTGAGCTACAA  
TbH9  
S P I S N M V S M A N N H M S M T N S G V S M T N T L S S M L  
GAAGGGCTTTGCTCCGGCGGGCGGCCGCCAGGCGGTGCAAACCGCGCGCAAAACGGGGTCCGGGCGATGAGCTCGCTGGGCAGCTCGCTG 1365  
CTTCCCGAAACGAGGCGCGCGCGGGTCCGGCACGTTTGGCGCGCGCTTTTGGCCAGGCCCGCTACTCGAGCGACCCGTCGAGCGAC  
TbH9  
K G F A P A A A A Q A V Q T A A Q N G V R A M S S L G S S L  
GGTTCTTCGGGTCTGGGCGGTGGGGTGGCCGCCAACTTGGGTGCGGCGGCTCGGTGCGTTCGTTGTCGGTGCCGACGGCTGGGCGCGG 1456  
CCAAGAAGCCCAGACCCGCCACCCACCGGCGGTGAACCCAGCCCGCGGAGCCAGCCAAGCAACAGCCACGGCTCCGGACCCGGCGCC  
TbH9  
G S S G L G G G V A A N L G R A A S V G S L S V P Q A W A A  
CCAACCAGGCAGTCACCCCGGCGGCGGGCGCTGCCGCTGACCAGCCTGACCAGCGCGCGGAAAGAGGGCCCGGGCAGATGCTGGGCGG 1547  
GGTTGGTCCGTCACTGGGGCCCGCGCGCGGCGGCGGACTGGTGGGACTGGTGGCGCGCTTTCTCCCGGGCCCGTCTACGACCCGCC  
TbH9  
A N Q A V T P A A R A L P L T S L T S A A E R G P G Q M L G G  
GCTGCCGGTGGGGCAGATGGGCGCCAGGGCCGGTGGTGGGCTCAGTGGTGTGCTGCGTGTTCGCGCGGACCTATGTGATGCCGATTCT 1638  
CGACGGCCACCCCGTCTACCCGCGGTCCCGGCCACCCAGTCAACACACGACGACAAGGCGGCGCTGGGATACACTACGGCGTAAGA  
TbH9  
L P V G Q M G A R A G G G L S G V L R V P P R P Y V M P H S

Fig. 11

Sheet 2 of 3

CCGGCAGCCGGCAAGCTTACTCAACTCCATTGCTTGCCAAATTGGCGGAGTTGGTCCGGGATTGCGGACATCATTTCCGGATG  
GGCGGTCCGCCGTTTGAATGAGTTGAGGAGTTTAACGAACGGTTTAACCGCCTCAACCAGCGCGCGGTAACGCCCTGTAGTAAAGCCTAC 1729  
TcH9 Hind3 hTCC1 (161-392)  
P A A G K L T Q L L K L L A K L A E L V A A A I A D I I S D  
TGGCGGACATCATCAAGGGCATCCTCGGAGAAGTGTGGGAGTTTCATCACAACCGCGCTCAACGGCCTGAAAGAGCTTTGGGACAAGCTCAC 1820  
ACCGCCTGTAGTAGTTCCCGTAGGAGCCTCTTACACCCCTCAAGTAGTGTGTTGGCGGAGTTGCCGGACTTTCTCGAAACCCTGTTCGAGTG  
hTCC1 (161-392)  
V A D I I K G I L G E V W E F I T N A L N G L K E L W O K L T  
GGGGTGGGTGACCGGACTGTTCTCTCGAGGGTGGTGAACCTGGAGTCTTCTTTGCGGGCGTCCCGGCTTGACCGGCGCGACCCAGCGGC 1911  
CCCCACCCACTGGCCTGACAAGAGAGCTCCACCAGCTTGACCTCAGGAAGAAACGCCCGCAGGGGCGGAACCTGGCCGCGCTGGTTCGGCG  
hTCC1 (161-392)  
G W V T G L F S R G W S N L E S F F A G V P G L T G A T S G  
TTGTCCGAAGTGACTGGCTTGTTCGGTGCGGCCGGTCTGTCCGCATCGTCCGGCTTGGCTCAGCGGATAGCCTGGCGAGCTCAGCCAGCT 2002  
AACAGCGTTCACTGACCGAACAAGCCACGCCCGCCAGACAGGCGTAGCAGCCCGAACCAGTGCAGCTATCGGACCGCTCGAGTCCGGTCTGA  
hTCC1 (161-392)  
L S Q V T G L F G A A G L S A S S G L A H A O S L A S S A S  
TGCCCGCCCTGGCCGGCATTGGGGGCGGCTCCGGTTTTGGGGGCTTCCGAGCCTGGCTCAGGTCCATGCCGCTCAACTCGGCAGGCGGT 2093  
ACGGGCGGGACCGGCCGTAACCCCGCCAGGCCAAAACCCCGAACGGCTCGGACCGAGTCCAGGTACGGCGGAGTTGAGCGCTCCGCGA  
hTCC1 (161-392)  
L P A L A G I G G G S G F G G L P S L A Q V H A A S T R Q A L  
ACGGCCCCGAGCTGATGGCCCGGTCCGGCGCCGCTGCCGAGCAGGTCCGGCGGGCAGTCCGAGCTGGTCTCCGCGCAGGGTTCCCAAGGTATG 2184  
TGCCGGGGCTCGACTACCGGGCCAGCCGCGGCGACGGCTCGTCCAGCCGCCCGTCAGCGTCGACCAGAGGCGCGTCCCAAGGGTTCCATAC  
hTCC1 (161-392)  
R P R A D G P V G A A A E Q V G G Q S C L V S A Q G S Q G M  
GGCGGACCCGTAGGCATGGGCGGCATGCACCCCTCTTCGGGGGCGTTCGAAAGGGACGACGACGAAGAAGTACTCGGAAGCGCGGCGGCGG 2275  
CCGCTGGGCATCCGTACCCGCCGTACGTGGGAGAAGCCCCCGCAGCTTTCCCTGCTGCTGCTTCTTCATGAGCCTTCCGCGCCGCGGCC  
hTCC1 (161-392)  
G G P V G M G G M H P S S G A S K G T T T K K Y S E G A A A  
GCACTGAAGACGCCGAGCGCGCGCCAGTCAAGCTGACCGGGCGGTGGGCAAAAGGTGCTGGTACGAAACGTCGTCTAACGGCGAATTC 2365  
CGTGACTTCTGCGGCTCGCGCGCGGTACGCTTCGACTGCGCCCGCCACCCGTTTTCCACGACCATGCTTTGCAGCAGATTGCCGCTTAAG  
hTCC1 (161-392) EcoRI  
G T E D A E R A P V E A D A G G G Q K V L V R N V V R R I

FIG. 11

Sheet 3 of 3

Enzymes: 3 of 515 enzymes (Filtered)

Samings: Linear, Certain Sites Standard Genetic Code

CATATGCATCACCATCACCATCAAGATGTGGCGGACATCAAGGGCATCTCGGAGAAGTGTGGGAGTTCAACAAAGCGGCTCAACGGCCTGAAAG  
GTATACGTAGTGGTAGTGGTAGTGTGTACACCGCCTGTAGTAGTTCCTGAGGAGCCTCTTCACACCCCTCAAGTAGTGTGTTGCGGAGTTGCGCGACTTTC  
Met / HIS TAG hTCC1 (184-392)  
H M H H H H H H D V A D I I K G I L G E V W E F I T N A L N G L K  
AGCTTTGGGACAAGCTCAGCGGGTGGGTGACCGGACTGTTCTCTCGAGGGTGGTGAACCTGGAGTCCTTCTTTGCGGGCTGCGCGGCTTACCGGGCG  
TCGAAACCTGTTCGAGTGCCTCCACCCACTGGCCTGACAAGAGAGCTCCACACAGCTTGGACCTCAGGAAGAAACGCCCGCAGGGGCGGAAGTGGCGCG  
hTCC1 (184-392)  
E L W D K L T G W V T G L F S R G W S N L E S F F A G V P G L T G A  
GACCAGCGGCTTGTGCAAGTGAAGTGGCTTGTTCGGTGGCGCGGTCTGTCCGCATCGTCGGGCTTGGCTCAGCGGATAGCCTGGCGAGCTCAGCCAGC  
CTGGTTCGCCGAACAGCGTTCAGTGAACGAACAAGCCAGCGCGGCGCAGACAGCGTAGCAGCCGAACCGAGTGGCGCTATCGGACCGCTCGAGTCGGTCA  
hTCC1 (184-392)  
T S G L S D V T G L F G A A G L S A S S G L A H A D S L A S S A S  
TTGCCCGCCTGGCGGCAATTGGGGGCGGGTCCGGTTTTGGGGGCTTGCCGAGCCTGGCTCAGGTCCATGCCGCTCAACTCGCGAGCGCTACGGCCCC  
AACGGGCGGGAGCGCGCTAACCCCGCGCCAGGCCAAAACCCCGGAACGGCTCGGACCGAGTCCAGGTACGGCGGAGTTGAGCGCTCCGCGATGCCGGGG  
hTCC1 (184-392)  
L P A L A G I G G G S G F G G L P S L A Q V H A A S T R Q A L R P  
GAGCTGATGGCCCGGTGGCGCGGCTGCCGAGCAGGTGGCGGGCAGTGGCAGCTGGTCTCCGCGCAGGGTTCCCAAGGTATGGCGGACCCGTAGGCAT  
CTCGACTACCGGGCCAGCGCGCGGACGGCTCGTCCAGCGCGCGTCCAGCTGACGAGAGGCGCGTCCCAAGGGTTCCATACCCGCTGGGCATCCGTA  
hTCC1 (184-392)  
R A D G P V G A A A E Q V G G Q S Q L V S A Q G S Q G M G G P V G M  
GGGCGGCATGCACCCCTCTTCGGGGGCGTGAAGGGGACGACGACGAAGAAGTACTCGGAAGGCGCGCGCGGGGCACTGAAGACGCGGAGCGCGGCCA  
CCGCGCTACGTGGGGAGAAGCCCCCGCAGCTTTCCTGCTGCTGCTTCTTCATGAGCCTTCCGCGCGCGCGCGGTGACTTCTGCGGCTCGCGCGCGGT  
hTCC1 (184-392)  
G G M H P S S G A S K G T T T K K Y S E G A A A G T E O A E R A P  
GTGGAAGCTGACGCGGGCGGTGGGCAAAAGGTGCTGGTACGAAACGTCGTGGAATTCATGGTGGATTTCGGGGCGTTACCACCGGAGATCAACTCCGCGA  
CAGCTTCGACTGCGCGCGCCACCCGTTTTCCACGACCATGCTTTGCAGCAGCTTAAGTACCACCTAAAGCCCCGCAATGGTGGCCTCTAGTTGAGGCGCT  
hTCC1 (184-392) EcoRI TbH9  
V E A D A G G G Q K V L V R N V V E F M V D F G A L P P E I N S A  
GGATGTACGCGGCGCGGTTTCGGCCTCGCTGGTGGCGCGGCTCAGATGTGGGACAGCGTGGCGAGTGACCTGTTTTCGGCCGCGTGGCGGTTTCAGTC  
CCTACATGCGGCGCGGCGCAAGCGGAGCGACACCGGCGCGGAGTCTACACCTGTGCGACCGCTCACTGGACAAAAGCGCGCAGCGCAAGTCAAG  
TbH9  
R M Y A G P G S A S L V A A A Q M W D S V A S D L F S A A S A F Q S  
GGTGGTCTGGGGTCTGACGGTGGGGTGGTGGATAGGTTTCGTGGCGGGTCTGATGGTGGCGGGCGCTCGCCGTATGTGGCGTGGATGAGCGTACCGCG  
CCACCAGACCCAGACTGCCACCCAGCACCTATCCAAGCAGCGCGCCAGACTACCACCGCGCGCGGAGCGGCATACACCGCACCTACTCGCAGTGGCGG  
TbH9  
V V W G L T V G S W I G S S A G L M V A A A S P Y V A W M S V T A

FIG. 12

Sheet 1 of 3

GGGCAGGCGGAGCTGACCGCGCGCCAGGTCCGCTGCGCGCGGCTACGAGACGGCGTATGGGCTGACGGTCCGCGCGGTGATGCGCCGAGAACC  
 CCGCTCGGCTCGACTGGCGCGGGTCCAGGCCAACGACGCGCGCGGATGCTCTGCCGATACCGGACTGCCACGCGCGCGGCTACTAGCGGCTCTTGG 1000

TbH9

G Q A E L T A A Q V R V A A A Y E T A Y G L T V P P P V I A E N  
 GTGCTGAAGTGTGATGATTCGTGATAGCGACCAACCTCTTGGGCAAAACACCGCGCGGATGCGCGTCAACGAGGCGGAATACGCGGAGATGTGGGCGCAAGA  
 CACGACTTGACTACTAAGACTATCGCTGGTTGGAGAACCCCGTTTTGTGGGCGCGCTAGCGCCAGTTGCTCCGGCTTATGCCGCTCTACACCGCGGTTCT 1100

TbH9

R A E L M I L I A T N L L G Q N T P A I A V N E A E Y G E M W A Q D  
 CGCGCGCGCGATGTTTGGCTACGCGCGCGCGACCGCGACGGCGACGGTGTGCTGCGGTTGAGGAGGCGCGCGGAGATGACCAGCGCGGGTGGGCTC  
 GCGCGCGCGCTACAAACCGATGCGCGCGCGCTGCGGCTGCGGCTGCGGCTGCAACGACGGCAAGCTCTCCGCGCGCTCTACTGGTGGCGGCCACCGGAG 1200

TbH9

A A A M F G Y A A A T A T A T A T L L P F E E A P E M T S A G G L  
 CTCGAGCAGGCGCGCGGTCGAGGAGGCTCCGACACCGCGCGCGGAACCGATTGATGAACAATGTGCCCCAGGCGCTGCAACAGCTGGCCCCAGCCCCA  
 GAGCTCGTCCGCGCGCGCGAGCTCTCCGAGGCTGTGGCGCGCGCGCTTGGTCAACTACTTGTACACGGGTCGCGGAGCTTGTGACCGGGTCGGGT 1300

TbH9

L E Q A A A V E E A S O T A A A N Q L M N N V P Q A L Q Q L A Q P  
 CGCAGGCGCACCGGCTTCTTCCAAGCTGGGTGGCTGTGGAAGACGGTCTCGCGCGATCGGTGCGCGATCAGCAACATGGTGTGATGGCCAACAACCA  
 GCGTCCCGTGGTGGGAAGAAGGTCGACCCACCGGACACCTTCTGCCAGAGCGCGGTAGCCAGCGGCTAGTCGTTGTACCACAGCTACCGGTTGTTGGT 1400

TbH9

T Q G T T P S S K L G G L W K T V S P H R S P I S N M V S M A N N H  
 CATGTCGATGACCAACTCGGGTGTGTCGATGACCAACACCTTGAGCTCGATGTTGAAGGCTTTGCTCCGCGCGCGCGCGCCAGGCGGTGCAAACCGCG  
 GTACAGCTACTGGTTGAGCCACACAGCTACTGGTTGTGGAACCTCGAGCTACAACCTTCCGAAACGAGGCGCGCGCGCGGGTCCGGCACGTTTGGCGC 1500

TbH9

M S M T N S G V S M T N T L S S M L K G F A P A A A A Q A V Q T A  
 GCGCAAAACGGGTCGCGCGGATGAGCTCGCTGGGCGAGCTCGCTGGGTTCTTCCGGTCTGGGCGGTGGGGTGGCGCGCAACTGGGTGCGGCGGCTCGG  
 CCGTTTTGCCCCAGGCGCGCTACTCGAGCGACCGCTCGAGCGACCCAAGAACGCCAGACCGCGCCACCGCGCGGTGAACCCAGCGCGCGGAGCC 1600

TbH9

A Q N G V R A M S S L G S S L G S S G L G G G V A A N L G R A A S  
 TCGGTTGTTGTGGTGCGCGCAGGCTGGGCGCGGCAACCAGGCAGTCACCCCGCGCGCGCGGCGCTGCCGCTGACCAGCCTGACCAGCGCGCGGA  
 AGCCAAGCAACAGCCACGGCTCGGACCGCGCGCGGTTGGTCCGTCAGTGGGCGCGCGCGCGCGCGGACGCGGACTGGTCCGACTGGTCCGCGCGCT 1700

TbH9

V G S L S V P Q A W A A A N Q A V T P A A R A L P L T S L T S A A E  
 AAGAGGCGCGCGGAGATGCTGGGCGGGTGGCGGTGGGCGAGATGGGCGCCAGGCGCGGTGGTGGGCTCAGTGGTGTGCTGCGTGTTCGCGCGGACCC  
 TTCTCCGCGGCGGTCTACGACCGCGCGGACGGCCACCGCTTACCGCGGTCCGCGCCACCGCGAGTCAACACAGCAGCACAAGGCGGCGCTGGG 1800

TbH9

R G P G Q M L G G L P V G Q M G A R A G G G L S G V L R V P P R P  
 TATGTGATGCCGATTCTCCGCGAGCGCGGATATCATGAGCAGAGCGTTTCATCATGATCCAACGATCAGTGCCATTGACGGCTTGTACGACCTTCTGG  
 ATACACTACGGCGTAAGAGGCGGTGGCGCGCTATAGTACTGCTCTCGCAAGTAGTAGCTAGGTTGCTAGTCACGGTAAGTGCCTGAACATGCTGGAAGACC 1900

TbH9 RV htCC1 (1-200)

Y V M P H S P A A G O I M S R A F I I O P T I S A I D G L Y O L L

FIG. 12

Sheet 2 of 3



GGATTGGAATACCCAACCAAGGGGTATCCTTCTCCTACTAGAGTACTTCGAAAAAGCCCTGGAGGAGCTTCAGCGTTTCGGGTGATGGCTG  
2000  
CCTAACCTTATGGGTTGGTTCCCCCATAGGAAATGAGGAGTATCTCATGAAGCTTTTTCGGGACCTCCTCGACCGTCTCGCAAGGCCCACTACCGAC  
hTCC1 (1-200)  
G I G I P N Q G G I L Y S S L E Y F E K A L E E L A A A F P G O G W  
GTTAGGTTTCGGCCGCGGACAAATACGCCGGCAAAAACCGCAACCACGTGAATTTTTTCAGGAACCTGGCAGACCTCGATCGTCAGCTCATCAGCCTGATC  
2100  
CAATCCAAGCCGGCGCCTGTTTATGCGGCCGTTTTTGGCCTTGGTGCACCTTAAAAAAGGTCTTTGACCGTCTGGAGCTAGCAGTCSAGTAGTCGGACTAG  
hTCC1 (1-200)  
L G S A A O K Y A G K N R N H V N F F Q E L A O L O R O L I S L I  
CACGACCAGGCCAACCGCGTCCAGACGACCCGCGACATCTGGAGGGCGCCAAGAAAGGTCTCGAGTTCTGTGCGCCCGGTGGCTGTGGACCTGACCTACA  
2200  
GTGCTGGTCCGGTTGCGCCAGGTCTGCTGGGCGCTGTAGGACCTCCCGCGGTTCTTTCCAGAGCTCAAGCACGCGGGCCACCGACACCTGGACTGGATGT  
hTCC1 (1-200)  
H O Q A N A V Q T T R O I L E G A K K G L E F V R P V A V O L T Y  
TCCCGGTCTGTCGGGCACGCCCTATCGGCCGCCCTTCCAGGCGCGTTTTTGCGCGGGCGCGATGGCCGTAGTGGGCGGCGCGCTTGCCTACTTGGTCTGTA  
2300  
AGGGCCAGCAGCCCGTGCGGGATAGCCGGCGGAAGGTCCGCGGCAAAAACGCCCGCGCTACCGGCATCACCGCGCGCGCAACGGATGAACCAGCACTT  
hTCC1 (1-200)  
I P V V G H A L S A A F Q A P F C A G A M A V V G G A L A Y L V V K  
AACGCTGATCAACGCGACTCAACTCCTCAAATTGCTTGCCAAATTGGCGGAGTTGGTCGCGGCGGCCATTGCGGACATCATTTCGGATGTGGCGGACATC  
2400  
TTGCGACTAGTTGCGCTGAGTTGAGGAGTTTAACGAACGGTTTAACCGCTCAACCAGCGCGCGGGTAACGCCTGTAGTAAAGCCTACACCGCCTGTAG  
hTCC1 (1-200)  
T L I N A T O L L K L L A K L A E L V A A A I A O I I S O V A O I  
ATCAAGGGCATCCTCGGAGAAGTGTGGGAGTTCATCTAAGATATC 2445  
TAGTTCCCGTAGGAGCCTCTTCACACCCCTCAAGTAGATTCTATAG  
hTCC1 (1-200) RV  
I K G I L G E V W E F I O I

FIG. 12

Sheet 3 of 3

Figure : Nucleotide sequence of MTb59

caccgactgcccgactgaacccgaactagtcagcacaaaaccgaagtaggaagacgaaaagctatggc  
 tgagttgacaatccccgctgatgacatccagagcgcgaatcgaagagtacgtaagctctttcaccgc  
 cgacaccagtagagaggaagtcgggtaccgtcgtcgatgccggggacggcatcgcacacgtcgaggg  
 tttgccatcgggtgatgacccaagagctgctcgaattcccgggcggaatcctcggcgtcgccctcaa  
 cctcgacgagcacagcgtcggcgcggtgatcctcgggtgacttcgagaacatcgaagaagggtcagca  
 ggtcaagcgcaccggcggaagtccttatcgggtcccggttgscgacgggtttttggggcggggtggttaa  
 cccgctcggccagccgatcgacggggcgcgagacgtcgactccgatactcggcgcgcgctggagct  
 ccaggcgccctcgggtggtgcaccggcaaggcgtgaaggagccgttgacagaccgggatcaaggcgat  
 tgacgcgatgacccccgatcggcccgggccagcgcagctgatcatcggcgaccgcaagaccggcaa  
 aaccgcccgtctgcgctcgacaccatcctcaaccagcggcagaactgggagtcgggtgatcccaagaa  
 gcagggtgcgctgtgtatagctggccatcgggcagaagggaactaccatcgccgcggtacgcccgcac  
 actggaagagggcggtgcatggactacaccaccatcgtcgcgggcgcgcggtcgaggtccgcggg  
 tttcaaatggcttgcgccgtacaccgggttcggcgatcgcccagcactggatgtacgagggcaagca  
 tgtgctgatcatcttcgacgacctgactaagcaggccgaggcataccggggcgatctcgctgctgct  
 ggcggctccgcccggcggtgaggcctaccccgcgatgtgttctatctgcattcgcggttttgga  
 gcgctgcgcccactgtccgacgatctcgggtggcggtcgctaacgggtctgccgatcatcgagac  
 caaggccaacgacatctcggcctacatcccgaccaacgtcatctcgatcaccgacgggcaatgttt  
 cctggaaaccgacctgttcaaccaggcggtccggccggccatcaacgtcgggtgtgctgggtgtcccg  
 agtcggcgggcgcgcgcgagatcaaggctatgaaagaggtcgcccgaagcctccgcttggaaccttc  
 gcaataaccgcgagctagaagctttcgccgctttcgcttctgatttggaacgcccgcacgaaggcgca  
 gttggagcgcgggcgcccggctggtcgagctgctcaagcagccgcaatcccagcccatgcccgttga  
 ggagcaagtgggtttcgatcttccctgggcaccggcggtcacctggactcgggtgcccgtcgaggacgt  
 ccggcggttcgaaaccgaattactggaccacatgcgggcctccgaagaagagattttgactgagat  
 ccgggacagccaaaagctcaccgaggaggccgcccgaagctcaccgagggtcatcaagaacttcaa  
 gaagggtctcgcgccaccgggtggcggtctgtgtggtgcccgaacatgtcgaggccctcgacga  
 ggataagctcgccaaggaagccgtgaagggtcaaaaagccggcgccgaagaagaagaatatagctaac  
 catggctgccacacttcgcgaactacgcggggcggtatcgctcggcagggtcgatcaaaaagatcac  
 caaggcccaggagctgatttgcgacatcgcgcatcgccagggcgagggctcggtcgagtcgctcg  
 gccctacgcttttgagatcacccggtgcttaccaccctggccgctgaagccgcactggaccatcc  
 gttgct

09688672-101000

Figure . : Amino acid sequence of MTb59

MAELTI PADDI QSAIEEYVSSFTADTSREEVGT VVDAGD GIAHVEGLPSVMTQELLEFPGGILGVA  
LNLDEHSVGAVILGDFENIEEGQQVKRTGEVLSVPVGDGFLGRV VNP LGCQPIDGRGDVSDTRRAL  
ELQAPS VVHRQGVKEPLQTGIKAIDAMTP IGRGQRQLI IGDRKTGKTAVCVDTILNQRQNWESCDP  
KKQVRCVYVAIGQKGT TIAAVRRTLEEGGAMDYTTIVAAAASESAGFKWLAPYTGSAIAQHWMYEG  
KHVLIIFDDLTKQAEAYRAISLLLRPPGREAYPGDV FYLHSRLLERC AKLSDDLGGGSLTGLPII  
ETKANDISAYIPTNVISITDGQC FLETDLFNQGV RPAINVGVSVSRVGGAAQIKAMKEVAGSLRLD  
LSQYRELEAF AAFASDLDAASKAQLERGARLV ELLKQPQS QPMPVEEQVVSIFLGTGGHLDSPVPE  
DVRRFETELLDHMRASEEEILTEIRDSQKLTEEAADKLTEVIKNFKKGFAATGGG SVVPDEHVEAL  
DEDKLAKEAVKVKKPAPKKKK

096867E 101000

Figure : Nucleotide sequence of MTb82

ccagcccccgccccgccccacgcccagggatgtgtggactgatggccaaagcgtcagagaccgaacgtt  
 cgggcccccgccaccaaccggcgagcggccagaccgacgtccgcgacgggttcgacccctgagca  
 cccaggcgggtgttccgccccgatctcggcgatgaggacaacttcccccatccgacgctcggccccg  
 acaccgagccgcaagaccggatggccaccaccagccgggtgcgcccggcggtcagacgggtgggcg  
 gggccttggtggaaatcccgcgggcgcccgatatcgatccgcttgaggccctgatgaccaaccgg  
 tggcgccggagtcgaagcgggttctgctggaactgtggacgtcccgtcggccgggtccgactcggaga  
 ccaaggaggacttcagaggggtgggtgtccctattgcggcagcccgatattcggttcctgcgcagctaa  
 atcccggggacatcgctcgccggccagtagcaggtcaaaggctgcacgcgcacggcgagactgggct  
 ggatctacctcgctctcgaccgcaatgtcaacggccgtccgggtgggtgctcaaggccctgggtgcatt  
 ccgggtgatgccgaagcgcaggcaatggcgatggccgaacgccaggttcctggccgaggtgggtgcacc  
 cgctcgatcgctgcagatcttcaactttgtcgagcacaccgacaggcacggggatccgggtcggctaca  
 tcgtgatggaatacgtcggcgggcaatcgctcaaacgcagcaagggtcagaaactgcccgctcgcg  
 aggccatcgctacctgctggagatcctgccggcgctgagctacctgcattccatcggccttggtct  
 acaacgacctgaagccggaaaacatcatgctgaccgaggaacagctcaagctgatcgacctgggcg  
 cggatccgcggtcaactcgcttcggctacctctacgggaccccaggcttccaggcgcccgagatcg  
 tgccgacccgggtccgacgggtggccaccgacatctacaccgtgggacgcacgctcgcggcgctcacgc  
 tggacctgcccaccgcaatggccgttatgtggatggggtacccgaagacgaccgggtgctgaaaa  
 cctacgactcttacggccgggtgctgcgcaggggccatcgaccccgatccgcggcaacgggttcacca  
 ccgcccgaagagatgtccgcgcaattgacggggcgtgttgccgggaggtgggtcgcccaggacaccgggg  
 tgccgcggccagggtatcaacgatcttcagtcaccagtcgggtcgacatttgagtggtgacctgctgg  
 tggcgcacaccgacgtgtatctggacggggcaggtgcacgcggagaaagctgaccgccaacgagatcg  
 tgaccgcgctgtcggtgcccgtggtcgatccgacccgacgtcgcagcttcgggtcctgcaggccacgg  
 tgctctcccagccgggtgcagaccctagactcgctgcgcgcggcccgccacgggtgcgctggacgcgc  
 acggcgtcgacttctccgagtcagtgagctgcccgtaatggaagtccgcgcgctgctggatctcg  
 gcgatgtggccaaggccaccgaaaaactcgacgatctggccgaacgcgttggttggtgatggcgat  
 tgggtctggtaaccggggcgtcgccgagctgctcacccggcgactatgactcggccaccaaacatttca  
 ccgaggtgctggataacctttcccgcgagctggcgcccaagctcgccctggccgcccaccgcccgaac  
 tagccggcaacaccgacgaacacaagttctatcagacgggtgtggagcaccacgacggcgctgatct  
 cggcgggtttcggactggccagagccgggtcggccgaagggtgatcgggtcggcgccgtgctgcacgc  
 tcgacgaggtaccgcccacttctcggcatttcaccacggcacggctgaccagcgcggtgactctgt  
 tgtccggccgggtcaacgagtgagtcaccgaggaacagatccgcgacccgcccgaagagtgagg  
 cgctgcccccgaccgaaccacgcgtgctgcagatccgcgccttggtgctgggtggcgcgctggact  
 ggctgaaggacaacaaggccagcaccaaccacatcctcggtttcccggttcaccagtcacgggctgc  
 ggctgggtgtcgaggcgtcactgcgcagcctggcccggttagctcccactcaacggcatcgctaca  
 cgctgggtggacatggccaacaaggtccggcccaccagcacgttctaagccgcccagtggtgaatcg

09655672-101000

Figure 1: Amino acid sequence of MTb82

MAKASETERSGPGTQPADAQTATSATVRPLSTQAVFRPDPFGDEDNFPHPPTLGPDTEPQDRMATTSR  
VRPPVRRLLGGCLVEIPRAPDIDPLEALMTNPVVPESKRFCWNCGRPVGRSDSETKGASEGWCPYCG  
SPYSFLPQLNPGLDIVAGQYEVKGCIAHGCLGWIYLAIDRNVNGRPVVLKGLVHSGDAEAQAMAMAE  
RQFLAEVVHPSIVQIFNFVEHTDRHGDPVGYIVMEYVGGQSLKRSKGQKLPVAEAIAYLLEILPAL  
SYLHSIGLVYNDLKPENIMLTTEEQLKLIDLGAVSRINSFGYLYGTPGFQAPEIVRTGPTVATDIYT  
VGRTLAALTLDLPTNRNGRYVDGLPEDDPVLKTYDSYGRLLRRAIDPDPRQRFTTAEEMSAQLTGVL  
REVVAQDTGVPRPGLSTIFSPSRSTFCVDLLVAHTDVYLDGQVHAEKLTANEIVTALSVPPLVDPTD  
VAASVLQATVLSQPVQTLDSLRAARHGALDADGVDFSESVELPLMEVRALLDLGDVAKATRKLLDDL  
AERVGWRWRLVWYRAVAELLTG DYDSATKHFTEVLDTFPGELAPKLALAATAELAGNTDEHKFYQT  
VWSTNDGVISAAFGGLARARSAEGDRVGAVRTLDEVPPTSRHFTTARLTSAVTLLSGRSTSEVTEEQ  
IRDAARRVEALPPTEPRVLQIRALVLGGALDWLKD NKASTNHILGFPFTSHGLRLGVEASLRSLAR  
VAPTQRHRYTLVDMANKVRPTSTF.

09688672-101000

17

Figure 17 Amino Acid Sequence of secreted DPPD

DPPDPHQPDMTKGYCPGGRWGFGLAVCDGEKYPDGSFWHQWMQTWFTGPPQFYFDCVSGGEPLP  
GPPPPGGCGCAIPSEQPNAP

09683672-101007